OM of: US-09-508-832-2 to: GenEmbl:* out_format : pfs

Date: Dec 11, 2001 1:45 AM

About: Results were produced by the GenCore software, version 4.5,
Copyright (c) 1993-2000 Compugen Ltd.

Command line parameters:
-MODEL=frame+_p2n.model -DEV=x1p
-Q=/cgn2_1/USPTO_spool/US09508832/runat_10122001_110349_29549/app_query.fasta_1.620
-DB=GenEmbl -QFMT=fastap -SUFFIX=rge -GAPOP=12.000 -GAPEXT=4.000
-MINMATCH=0.100 -LOOPCL=0.000 -LOOPEXT=0.000 -QGAPOP=4.500
-QGAPEXT=0.050 -XGAPOP=10.000 -XGAPEXT=0.500 -FGAPOP=6.000
-FGAPEXT=7.000 -YGAPOP=10.000 -YGAPEXT=0.500 -DELOP=6.000
-DELEXT=7.000 -START=1 -MATRIX=blosum62 -TRANS=human40.cdi
-LIST=45 -DOCALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0
-ALIGN=15 -MODE=LOCAL -OUTFMT=pfs -NORM=ext -MINLEN=0
-MAXLEN=2000000000 -USER=US09508832_@CGN1_1_0 -NCPU=6 -ICPU=3
-LONGLOG -NO_XLPXY -WAIT -THREADS=1

Search information block:
Query: US-09-508-832-2
Query length: 110
Database: GenEmbl:*
Database sequences: 1472140
Database length: -341344837
Search time (sec): 2485.920000

score_list:

Sequence	Strd	Orig	ZScore	EScore	Len	Documentation
gb_pat:AX031279	+	574.00	959.38	3.7e-45	332	AX031279 Sequence 1 from Patent
gb_un:AX031305	+	574.00	959.38	3.7e-45	332	AX031305 Sequence 1 from Patent
gb_ro:AF032461	+	574.00	959.35	3.7e-45	333	AF032461 Mus musculus BimS mRNA
gb_ro:AF065432	+	556.00	929.37	1.7e-43	333	AF065432 Rattus norvegicus Bcl-
gb_pat:AX031281	+	549.00	915.94	9.7e-43	422	AX031281 Sequence 3 from Patent
gb_un:AX031307	+	549.00	915.94	9.7e-43	422	AX031307 Sequence 3 from Patent
gb_ro:AF032460	+	549.00	915.92	9.7e-43	423	AF032460 Mus musculus BimL mRNA
gb_ro:AF136927	+	534.00	890.94	2.4e-41	423	AF136927 Rattus norvegicus Bcl-
gb_pat:AX031283	+	521.00	866.79	5.3e-40	590	AX031283 Sequence 5 from Patent
gb_un:AX031309	+	521.00	866.79	5.3e-40	590	AX031309 Sequence 5 from Patent
gb_ro:AF032459	+	521.00	866.78	5.3e-40	591	AF032459 Mus musculus BimEL mRN
gb_ro:AF065433	+	506.00	841.79	1.3e-38	591	AF065433 Rattus norvegicus Bcl-
gb_pat:AX031285	+	436.00	727.84	2.9e-32	416	AX031285 Sequence 7 from Patent
gb_un:AX031311	+	436.00	727.84	2.9e-32	416	AX031311 Sequence 7 from Patent
gb_pr:AF032458	+	436.00	727.82	2.9e-32	417	AF032458 Homo sapiens BimL mRNA
gb_pat:AX031287	+	406.00	675.18	2.5e-29	596	AX031287 Sequence 9 from Patent
gb_un:AX031313	+	406.00	675.18	2.5e-29	596	AX031313 Sequence 9 from Patent
gb_pr:AF032457	+	406.00	675.16	2.5e-29	597	AF032457 Homo sapiens BimEL mRN
gb_ro:AF065431	+	353.00	592.51	1.0e-24	282	AF065431 Rattus norvegicus Bcl-
gb_htg:AC013332	-	210.00	305.89	9.2e-09	180569	AC013332 Homo sapiens chromo
gb_htg:AC013332	+	132.00	175.97	0.1589	180569	AC013332 Homo sapiens chromo
gb_htg:AC046192	-	129.00	172.37	0.2524	150035	AC046192 Homo sapiens chromo
gb_htg:AC046192	+	128.00	170.70	0.3125	150035	AC046192 Homo sapiens chromo
gb_ba:SCD63A	-	93.00	123.40	134.69	34611	AL356832 Streptomyces coelico
gb_pr:HSU63836	+	89.00	144.55	8.95	849	U63836 Human high molecular wei
gb_pl:ACBTUB	+	88.50	136.55	24.94	2206	X72789 A.chrysogenum gene for
gb_htg:AC091680	-	88.50	104.91	1.4e+03	149967	AC091680 Oryza sativa chromo
gb_htg:AL161774	+	86.50	100.16	2.7e+03	181386	AL161774 Homo sapiens chromo
gb_pr:HSMUC5BEX	+	86.00	120.54	194.50	10713	Z72496 H.sapiens MUC5B gene (
gb_htg:AC061979	-	86.00	99.61	2.8e+03	174511	AC061979 Homo sapiens chromo
gb_pat:A39632	+	84.50	129.89	58.62	2206	A39632 Sequence 17 from Patent
gb_pat:E05189	+	84.50	126.55	89.99	3445	E05189 DNA sequence coding for
gb_in:AC007581	-	84.50	97.93	3.5e+03	156508	AC007581 Drosophila melanoga
gb_in:AC007925	-	84.50	97.31	3.8e+03	170089	AC007925 Drosophila melanoga
gb_htg:AC025106	-	84.50	96.44	4.3e+03	190982	AC025106 Homo sapiens chromo
gb_in:AE003465	-	84.50	92.37	7.2e+03	328500	AE003465 Drosophila melanoga
gb_ro:AF173359	-	84.00	134.39	32.91	1083	AF173359 Mus musculus Nesp-as
gb_ro:AF175305	+	84.00	134.20	33.72	1111	AF175305 Mus musculus Nesp mRN
gb_ro:MMU010163	+	84.00	133.42	37.28	1233	AJ010163 Mus musculus mRNA for
gb_ro:AF107848	+	84.00	128.23	72.52	2463	AF107848 Mus musculus neuroend
gb_ro:MMU245401	+	84.00	124.94	110.63	3821	AJ245401 Mus musculus Nesp gen
gb_ro:MMU251761	+	84.00	111.84	593.78	21927	AJ251761 Mus musculus Nesp
gb_htg:AL450323	+	84.00	99.74	2.8e+03	110015	AL450323 Mus musculus chr
gb_htg:AL593857	+	84.00	95.03	5.1e+03	206137	AL593857 Mus musculus chr
gb_pl:AF110419	-	83.50	127.72	77.48	2361	AF110419 Lepocinclis ovum s

seq_name: gb_pat:AX031279

seq_documentation_block:
LOCUS AX031279 332 bp DNA PAT 20-SEP-2000
DEFINITION Sequence 1 from Patent WO9914321.
ACCESSION AX031279
VERSION AX031279.1 GI:10278610
KEYWORDS .
SOURCE unidentified.
ORGANISM unidentified
unclassified.
REFERENCE 1 (bases 1 to 332)
AUTHORS O'Reilly,L., Puthalakath,H., Adams,J., O'Connor,L., Cory,S.,
Huang,D.C. and Strasser,A.
TITLE Novel therapeutic molecules
JOURNAL Patent: WO 9914321-A 1 25-MAR-1999;
INST MEDICAL W & E HALL (AU) ; PUTHALAKATH HAMSA (AU) ; REILLY
LORRAINE O (AU) ; ADAMS JERRY (AU) ; CONNOR LIAM O (AU) ; CORY
SUZANNE (AU) ; HUANG DAVID C S (AU) ; STRASSER ANDREAS (AU)
FEATURES Location/Qualifiers
source 1. .332
/organism="unidentified"
/db_xref="taxon:32644"
CDS 1. .>330
/note="unnamed protein product"
/codon_start=1
/protein_id="CAC09653.1"
/db_xref="GI:10278611"
/translation="MAKQPSDVSSECDREGGQLQPAERPPQLRPGAPTSLQTEPQASI
RQSQEEPEDLRPEIRIAQELRRIGDEFNETYTRRVFANDYREAEDHPQMVILQLLRFI
FRLVWRRH"
BASE COUNT 87 a 85 c 91 g 69 t
ORIGIN

alignment_scores:
Quality: 574.00 Length: 110
Ratio: 5.218 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
US-09-508-832-2 x AX031279 ..

Align seg 1/1 to: AX031279 from: 1 to: 332

```
      1 MetAlaLysGlnProSerAspValSerSerGluCysAspArgGluGlyGl 17
        |||||||||||||||||||||||||||||||||||||||||||||||||||
      1 ATGGCCAAGCAACCTTCTGATGTAAGTTCTGAGTGTGACAGAGAAGGTGG 50

     17 yGlnLeuGlnProAlaGluArgProProGlnLeuArgProGlyAlaProT 34
        |||||||||||||||||||||||||||||||||||||||||||||||||||
     51 ACAATTGCAGCCTGCTGAGAGGCCTCCCCAGCTCAGGCCTGGGGCCCCTA 100

     34 hrSerLeuGlnThrGluProGlnAlaSerIleArgGlnSerGlnGluGlu 50
        |||||||||||||||||||||||||||||||||||||||||||||||||||
    101 CCTCCCTACAGACAGAACCGCAAGCTTCCATACGACAGTCTCAGGAGGAA 150

     51 ProGluAspLeuArgProGluIleArgIleAlaGlnGluLeuArgArgIl 67
        |||||||||||||||||||||||||||||||||||||||||||||||||||
    151 CCTGAAGATCTGCGCCCGGAGATACGGATTGCACAGGAGCTGCGGCGGAT 200

     67 eGlyAspGluPheAsnGluThrTyrThrArgArgValPheAlaAsnAspT 84
        |||||||||||||||||||||||||||||||||||||||||||||||||||
    201 CGGAGACGAGTTCAACGAAACTTACACAAGGAGGGTGTTTGCAAATGATT 250

     84 yrArgGluAlaGluAspHisProGlnMetValIleLeuGlnLeuLeuArg 100
        |||||||||||||||||||||||||||||||||||||||||||||||||||
    251 ACCGCGAGGCTGAAGACCACCCTCAAATGGTTATCTTACAACTGTTACGC 300
```

```
     101 PheIlePheArgLeuValTrpArgArgHis 110
         ||||||||||||||||||||||||||||||
     301 TTTATCTTCCGTCTGGTATGGAGAAGGCAT 330

seq_name: gb_un:AX031305

seq_documentation_block:
LOCUS       AX031305      332 bp    DNA             UNA       20-SEP-2000
DEFINITION  Sequence 1 from Patent WO9914321.
ACCESSION   AX031305
VERSION     AX031305.1  GI:10278633
KEYWORDS    .
SOURCE      unidentified.
  ORGANISM  unidentified
            unclassified.
REFERENCE   1  (bases 1 to 332)
  AUTHORS   O'Reilly,L., Puthalakath,H., Adams,J., O'Connor,L., Cory,S.,
            Huang,D.C. and Strasser,A.
  TITLE     Novel therapeutic molecules
  JOURNAL   Patent: WO 9914321-A 25-MAR-1999;
            INST MEDICAL W & E HALL (AU) ; PUTHALAKATH HAMSA (AU) ; REILLY
            LORRAINE O (AU) ; ADAMS JERRY (AU) ; CONNOR LIAM O (AU) ; CORY
            SUZANNE (AU) ; HUANG DAVID C S (AU) ; STRASSER ANDREAS (AU)
FEATURES             Location/Qualifiers
     source          1. .332
                     /organism="unidentified"
                     /db_xref="taxon:32644"
     CDS             1. .>330
                     /note="unnamed protein product"
                     /codon_start=1
                     /protein_id="CAC09658.1"
                     /db_xref="GI:10278634"
                     /translation="MAKQPSDVSSECDREGGQLQPAERPPQLRPGAPTSLQTEPQASI
                     RQSQEEPEDLRPEIRIAQELRRIGDEFNETYTRRVFANDYREAEDHPQMVILQLLRFI
                     FRLVWRRH"
BASE COUNT       87 a     85 c     91 g     69 t
ORIGIN


alignment_scores:
            Quality:  574.00            Length:    110
              Ratio:  5.218               Gaps:      0
 Percent Similarity: 100.000   Percent Identity: 100.000

alignment_block:
 US-09-508-832-2 x AX031305   ..

 Align seg 1/1  to: AX031305  from: 1  to: 332

       1 MetAlaLysGlnProSerAspValSerSerGluCysAspArgGluGlyGl 17
         ||||||||||||||||||||||||||||||||||||||||||||||||||
       1 ATGGCCAAGCAACCTTCTGATGTAAGTTCTGAGTGTGACAGAGAAGGTGG 50

      17 yGlnLeuGlnProAlaGluArgProProGlnLeuArgProGlyAlaProT 34
         ||||||||||||||||||||||||||||||||||||||||||||||||||
      51 ACAATTGCAGCCTGCTGAGAGGCCTCCCCAGCTCAGGCCTGGGGCCCCTA 100

      34 hrSerLeuGlnThrGluProGlnAlaSerIleArgGlnSerGlnGluGlu 50
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     101 CCTCCCTACAGACAGAACCGCAAGCTTCCATACGACAGTCTCAGGAGGAA 150

      51 ProGluAspLeuArgProGluIleArgIleAlaGlnGluLeuArgArgIl 67
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     151 CCTGAAGATCTGCGCCCGGAGATACGGATTGCACAGGAGCTGCGGCGGAT 200

      67 eGlyAspGluPheAsnGluThrTyrThrArgArgValPheAlaAsnAspT 84
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     201 CGGAGACGAGTTCAACGAAACTTACACAAGGAGGGTGTTTGCAAATGATT 250

      84 yrArgGluAlaGluAspHisProGlnMetValIleLeuGlnLeuLeuArg 100
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     251 ACCGCGAGGCTGAAGACCACCCTCAAATGGTTATCTTACAACTGTTACGC 300
```

```
     101 PheIlePheArgLeuValTrpArgArgHis 110
         ||||||||||||||||||||||||||||||
     301 TTTATCTTCCGTCTGGTATGGAGAAGGCAT 330

seq_name: gb_ro:AF032461

seq_documentation_block:
LOCUS       AF032461      333 bp    mRNA            ROD       19-FEB-1998
DEFINITION  Mus musculus BimS mRNA, complete cds.
ACCESSION   AF032461
VERSION     AF032461.1  GI:2895503
KEYWORDS    .
SOURCE      house mouse.
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 333)
  AUTHORS   O'Connor,L., Strasser,A., O'Reilly,L.A., Hausmann,G., Adams,J.M.,
            Cory,S. and Huang,D.C.
  TITLE     Bim: a novel member of the Bcl-2 family that promotes apoptosis
  JOURNAL   EMBO J. 17 (2), 384-395 (1998)
  MEDLINE   98094360
   PUBMED   9430630
REFERENCE   2  (bases 1 to 333)
  AUTHORS   O'Connor,L., Strasser,A., O'Reilly,L.A., Hausmann,G., Adams,J.M.,
            Cory,S. and Huang,D.C.S.
  TITLE     Direct Submission
  JOURNAL   Submitted (03-NOV-1997) Molecular Genetics of Cancer, The Walter &
            Eliza Hall Institute of Medical Research, PO Royal Melbourne
            Hospital, Parkville, Victoria 3050, Australia
FEATURES             Location/Qualifiers
     source          1. .333
                     /organism="Mus musculus"
                     /db_xref="taxon:10090"
     CDS             1. .333
                     /note="pro-apoptotic BH3-containing Bcl-2 family member"
                     /codon_start=1
                     /product="BimS"
                     /protein_id="AAC40031.1"
                     /db_xref="GI:2895504"
                     /translation="MAKQPSDVSSECDREGGQLQPAERPPQLRPGAPTSLQTEPQASI
                     RQSQEEPEDLRPEIRIAQELRRIGDEFNETYTRRVFANDYREAEDHPQMVILQLLRFI
                     FRLVWRRH"
BASE COUNT       88 a     85 c     91 g     69 t
ORIGIN


alignment_scores:
            Quality:  574.00            Length:    110
              Ratio:  5.218               Gaps:      0
 Percent Similarity: 100.000   Percent Identity: 100.000

alignment_block:
 US-09-508-832-2 x AF032461   ..

 Align seg 1/1  to: AF032461  from: 1  to: 333

       1 MetAlaLysGlnProSerAspValSerSerGluCysAspArgGluGlyGl 17
         ||||||||||||||||||||||||||||||||||||||||||||||||||
       1 ATGGCCAAGCAACCTTCTGATGTAAGTTCTGAGTGTGACAGAGAAGGTGG 50

      17 yGlnLeuGlnProAlaGluArgProProGlnLeuArgProGlyAlaProT 34
         ||||||||||||||||||||||||||||||||||||||||||||||||||
      51 ACAATTGCAGCCTGCTGAGAGGCCTCCCCAGCTCAGGCCTGGGGCCCCTA 100

      34 hrSerLeuGlnThrGluProGlnAlaSerIleArgGlnSerGlnGluGlu 50
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     101 CCTCCCTACAGACAGAACCGCAAGCTTCCATACGACAGTCTCAGGAGGAA 150

      51 ProGluAspLeuArgProGluIleArgIleAlaGlnGluLeuArgArgIl 67
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     151 CCTGAAGATCTGCGCCCGGAGATACGGATTGCACAGGAGCTGCGGCGGAT 200
```

```
     67 eGlyAspGluPheAsnGluThrTyrThrArgArgValPheAlaAsnAspT 84
        ||||||||||||||||||||||||||||||||||||||||||||||||||
    201 CGGAGACGAGTTCAACGAAACTTACACAAGGAGGGTGTTTGCAAATGATT 250

     84 yrArgGluAlaGluAspHisProGlnMetValIleLeuGlnLeuLeuArg 100
        ||||||||||||||||||||||||||||||||||||||||||||||||||
    251 ACCGCGAGGCTGAAGACCACCCTCAAATGGTTATCTTACAACTGTTACGC 300

    101 PheIlePheArgLeuValTrpArgArgHis 110
        ||||||||||||||||||||||||||||||
    301 TTTATCTTCCGTCTGGTATGGAGAAGGCAT 330

seq_name: gb_ro:AF065432

seq_documentation_block:
LOCUS       AF065432      333 bp    mRNA            ROD       11-MAR-1999
DEFINITION  Rattus norvegicus Bcl-2 related ovarian death gene product BOD-M
            mRNA, complete cds.
ACCESSION   AF065432
VERSION     AF065432.1  GI:3228567
KEYWORDS    .
SOURCE      Norway rat.
  ORGANISM  Rattus norvegicus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
            Rattus.
REFERENCE   1  (bases 1 to 333)
  AUTHORS   Hsu,S.Y., Lin,P. and Hsueh,A.J.
  TITLE     BOD (Bcl-2-related ovarian death gene) is an ovarian BH3
            domain-containing proapoptotic Bcl-2 protein capable of
            dimerization with diverse antiapoptotic Bcl-2 members
  JOURNAL   Mol. Endocrinol. 12 (9), 1432-1440 (1998)
  MEDLINE   98400436
REFERENCE   2  (bases 1 to 333)
  AUTHORS   Hsu,S.Y. and Hsueh,A.J.W.
  TITLE     Direct Submission
  JOURNAL   Submitted (15-MAY-1998) GYN/OB, Stanford University, MSOB S385,
            Stanford, CA 94305, USA
FEATURES             Location/Qualifiers
     source          1. .333
                     /organism="Rattus norvegicus"
                     /db_xref="taxon:10116"
     CDS             1. .333
                     /codon_start=1
                     /product="Bcl-2 related ovarian death gene product BOD-M"
                     /protein_id="AAC23594.1"
                     /db_xref="GI:3228568"
                     /translation="MAKQPSDVNSECDREGGQLQPAERPPQLRPGAPTSLQTESQASI
                     RQSQEDPEDLRPEIRIAQELRRIGDEFNETYTRRAFANDYREAEDHPQMVILQLLRFI
                     FRLVWRRH"
BASE COUNT      88 a     86 c     94 g     65 t
ORIGIN


alignment_scores:
            Quality:  556.00             Length:    110
              Ratio:  5.101                Gaps:      0
 Percent Similarity:  99.091   Percent Identity: 96.364

alignment_block:
 US-09-508-832-2 x AF065432   ..

 Align seg 1/1  to: AF065432  from: 1  to: 333

      1 MetAlaLysGlnProSerAspValSerSerGluCysAspArgGluGlyGl 17
        ||||||||||||||||||||||:::||||||||||||||||||||||||||
      1 ATGGCCAAGCAACCTTCTGATGTAAATTCTGAGTGTGACAGAGAAGGTGG 50

     17 yGlnLeuGlnProAlaGluArgProProGlnLeuArgProGlyAlaProT 34
        ||||||||||||||||||||||||||||||||||||||||||||||||||
     51 ACAATTGCAGCCTGCTGAGAGGCCTCCCCAGCTCAGGCCTGGGGCCCCTA 100

     34 hrSerLeuGlnThrGluProGlnAlaSerIleArgGlnSerGlnGluGlu 50
        ||||||||||||||||||   ||||||||||||||||||||||||||:::
    101 CCTCCCTACAGACAGAATCGCAAGCTTCCATAAGGCAGTCTCAGGAGGAC 150

     51 ProGluAspLeuArgProGluIleArgIleAlaGlnGluLeuArgArgIl 67
        ||||||||||||||||||||||||||||||||||||||||||||||||||
    151 CCTGAAGATTTGCGCCCAGAGATACGGATCGCACAGGAGCTGCGGCGGAT 200

     67 eGlyAspGluPheAsnGluThrTyrThrArgArgValPheAlaAsnAspT 84
        ||||||||||||||||||||||||||||||||||:::||||||||||||||
    201 CGGAGACGAGTTCAATGAGACTTACACGAGGAGGGCGTTTGCAAACGATT 250

     84 yrArgGluAlaGluAspHisProGlnMetValIleLeuGlnLeuLeuArg 100
        ||||||||||||||||||||||||||||||||||||||||||||||||||
    251 ACCGAGAGGCGGAAGACCACCCGCAAATGGTTATCTTACAACTGTTACGA 300

    101 PheIlePheArgLeuValTrpArgArgHis 110
        ||||||||||||||||||||||||||||||
    301 TTCATCTTCCGTCTGGTCTGGAGAAGGCAC 330

seq_name: gb_pat:AX031281

seq_documentation_block:
LOCUS       AX031281      422 bp    DNA             PAT       20-SEP-2000
DEFINITION  Sequence 3 from Patent WO9914321.
ACCESSION   AX031281
VERSION     AX031281.1  GI:10278612
KEYWORDS    .
SOURCE      unidentified.
  ORGANISM  unidentified
            unclassified.
REFERENCE   1  (bases 1 to 422)
  AUTHORS   O'Reilly,L., Puthalakath,H., Adams,J., O'Connor,L., Cory,S.,
            Huang,D.C. and Strasser,A.
  TITLE     Novel therapeutic molecules
  JOURNAL   Patent: WO 9914321-A 3 25-MAR-1999;
            INST MEDICAL W & E HALL (AU) ; PUTHALAKATH HAMSA (AU) ; REILLY
            LORRAINE O (AU) ; ADAMS JERRY (AU) ; CONNOR LIAM O (AU) ; CORY
            SUZANNE (AU) ; HUANG DAVID C S (AU) ; STRASSER ANDREAS (AU)
FEATURES             Location/Qualifiers
     source          1. .422
                     /organism="unidentified"
                     /db_xref="taxon:32644"
     CDS             1. .>420
                     /note="unnamed protein product"
                     /codon_start=1
                     /protein_id="CAC09654.1"
                     /db_xref="GI:10278613"
                     /translation="MAKQPSDVSSECDREGGQLQPAERPPQLRPGAPTSLQTEPQDRS
                     PAPMSCDKSTQTPSPPCQAFNHYLSAMASIRQSQEEPEDLRPEIRIAQELRRIGDEFN
                     ETYTRRVFANDYREAEDHPQMVILQLLRFIFRLVWRRH"
BASE COUNT     112 a    116 c    109 g     85 t
ORIGIN


alignment_scores:
            Quality:  549.00             Length:    140
              Ratio:  4.991                Gaps:      1
 Percent Similarity:  78.571   Percent Identity: 78.571

alignment_block:
 US-09-508-832-2 x AX031281   ..

 Align seg 1/1  to: AX031281  from: 1  to: 422

      1 MetAlaLysGlnProSerAspValSerSerGluCysAspArgGluGlyGl 17
        ||||||||||||||||||||||||||||||||||||||||||||||||||||
      1 ATGGCCAAGCAACCTTCTGATGTAAGTTCTGAGTGTGACAGAGAAGGTGG 50

     17 yGlnLeuGlnProAlaGluArgProProGlnLeuArgProGlyAlaProT 34
        ||||||||||||||||||||||||||||||||||||||||||||||||||
     51 ACAATTGCAGCCTGCTGAGAGGCCTCCCCAGCTCAGGCCTGGGGCCCCTA 100
```

```
     34 hrSerLeuGlnThrGluProGln..... 41
        ||||||||||||||||||||||
    101 CCTCCCTACAGACAGAACCGCAAGACAGGAGCCCGGCACCCATGAGTTGT 150

     41 ..... 41

    151 GACAAGTCAACACAAACCCCAAGTCCTCCTTGCCAGGCCTTCAACCACTA 200

     42 .....AlaSerIleArgGlnSerGlnGluGluProGluAspL 54
                ||||||||||||||||||||||||||||||||||||
    201 TCTCAGTGCAATGGCTTCCATACGACAGTCTCAGGAGGAACCTGAAGATC 250

     54 euArgProGluIleArgIleAlaGlnGluLeuArgArgIleGlyAspGlu 70
        ||||||||||||||||||||||||||||||||||||||||||||||||||
    251 TGCGCCCGGAGATACGGATTGCACAGGAGCTGCGGCGGATCGGAGACGAG 300

     71 PheAsnGluThrTyrThrArgArgValPheAlaAsnAspTyrArgGluAl 87
        ||||||||||||||||||||||||||||||||||||||||||||||||||
    301 TTCAACGAAACTTACACAAGGAGGGTGTTTGCAAATGATTACCGCGAGGC 350

     87 aGluAspHisProGlnMetValIleLeuGlnLeuLeuArgPheIlePheA 104
        ||||||||||||||||||||||||||||||||||||||||||||||||||
    351 TGAAGACCACCCTCAAATGGTTATCTTACAACTGTTACGCTTTATCTTCC 400

    104 rgLeuValTrpArgArgHis 110
        ||||||||||||||||||||
    401 GTCTGGTATGGAGAAGGCAT 420
```

```
seq_name: gb_un:AX031307

seq_documentation_block:
LOCUS       AX031307      422 bp    DNA             UNA       20-SEP-2000
DEFINITION  Sequence 3 from Patent WO9914321.
ACCESSION   AX031307
VERSION     AX031307.1  GI:10278635
KEYWORDS    .
SOURCE      unidentified.
  ORGANISM  unidentified
            unclassified.
REFERENCE   1  (bases 1 to 422)
  AUTHORS   O'Reilly,L., Puthalakath,H., Adams,J., O'Connor,L., Cory,S.,
            Huang,D.C. and Strasser,A.
  TITLE     Novel therapeutic molecules
  JOURNAL   Patent: WO 9914321-A 25-MAR-1999;
            INST MEDICAL W & E HALL (AU) ; PUTHALAKATH HAMSA (AU) ; REILLY
            LORRAINE O (AU) ; ADAMS JERRY (AU) ; CONNOR LIAM O (AU) ; CORY
            SUZANNE (AU) ; HUANG DAVID C S (AU) ; STRASSER ANDREAS (AU)
FEATURES             Location/Qualifiers
     source          1. .422
                     /organism="unidentified"
                     /db_xref="taxon:32644"
     CDS             1. .>420
                     /note="unnamed protein product"
                     /codon_start=1
                     /protein_id="CAC09659.1"
                     /db_xref="GI:10278636"
                     /translation="MAKQPSDVSSECDREGGQLQPAERPPQLRPGAPTSLQTEPQDRS
                     PAPMSCDKSTQTPSPPCQAFNHYLSAMASIRQSQEEPEDLRPEIRIAQELRRIGDEFN
                     ETYTRRVFANDYREAEDHPQMVILQLLRFIFRLVWRRH"
BASE COUNT      112 a    116 c    109 g     85 t
ORIGIN


alignment_scores:
            Quality:  549.00            Length:    140
              Ratio:  4.991               Gaps:      1
 Percent Similarity:  78.571   Percent Identity: 78.571

alignment_block:
 US-09-508-832-2 x AX031307   ..

 Align seg 1/1  to: AX031307  from: 1  to: 422
```

```
      1 MetAlaLysGlnProSerAspValSerSerGluCysAspArgGluGlyGl 17
        ||||||||||||||||||||||||||||||||||||||||||||||||||
      1 ATGGCCAAGCAACCTTCTGATGTAAGTTCTGAGTGTGACAGAGAAGGTGG 50

     17 yGlnLeuGlnProAlaGluArgProProGlnLeuArgProGlyAlaProT 34
        ||||||||||||||||||||||||||||||||||||||||||||||||||
     51 ACAATTGCAGCCTGCTGAGAGGCCTCCCCAGCTCAGGCCTGGGGCCCCTA 100

     34 hrSerLeuGlnThrGluProGln..... 41
        ||||||||||||||||||||||
    101 CCTCCCTACAGACAGAACCGCAAGACAGGAGCCCGGCACCCATGAGTTGT 150

     41 ..... 41

    151 GACAAGTCAACACAAACCCCAAGTCCTCCTTGCCAGGCCTTCAACCACTA 200

     42 .....AlaSerIleArgGlnSerGlnGluGluProGluAspL 54
                ||||||||||||||||||||||||||||||||||||
    201 TCTCAGTGCAATGGCTTCCATACGACAGTCTCAGGAGGAACCTGAAGATC 250

     54 euArgProGluIleArgIleAlaGlnGluLeuArgArgIleGlyAspGlu 70
        ||||||||||||||||||||||||||||||||||||||||||||||||||
    251 TGCGCCCGGAGATACGGATTGCACAGGAGCTGCGGCGGATCGGAGACGAG 300

     71 PheAsnGluThrTyrThrArgArgValPheAlaAsnAspTyrArgGluAl 87
        ||||||||||||||||||||||||||||||||||||||||||||||||||
    301 TTCAACGAAACTTACACAAGGAGGGTGTTTGCAAATGATTACCGCGAGGC 350

     87 aGluAspHisProGlnMetValIleLeuGlnLeuLeuArgPheIlePheA 104
        ||||||||||||||||||||||||||||||||||||||||||||||||||
    351 TGAAGACCACCCTCAAATGGTTATCTTACAACTGTTACGCTTTATCTTCC 400

    104 rgLeuValTrpArgArgHis 110
        ||||||||||||||||||||
    401 GTCTGGTATGGAGAAGGCAT 420
```

```
seq_name: gb_ro:AF032460

seq_documentation_block:
LOCUS       AF032460      423 bp    mRNA            ROD       19-FEB-1998
DEFINITION  Mus musculus BimL mRNA, complete cds.
ACCESSION   AF032460
VERSION     AF032460.1  GI:2895501
KEYWORDS    .
SOURCE      house mouse.
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 423)
  AUTHORS   O'Connor,L., Strasser,A., O'Reilly,L.A., Hausmann,G., Adams,J.M.,
            Cory,S. and Huang,D.C.
  TITLE     Bim: a novel member of the Bcl-2 family that promotes apoptosis
  JOURNAL   EMBO J. 17 (2), 384-395 (1998)
  MEDLINE   98094360
   PUBMED   9430630
REFERENCE   2  (bases 1 to 423)
  AUTHORS   O'Connor,L., Strasser,A., O'Reilly,L.A., Hausmann,G., Adams,J.M.,
            Cory,S. and Huang,D.C.S.
  TITLE     Direct Submission
  JOURNAL   Submitted (03-NOV-1997) Molecular Genetics of Cancer, The Walter &
            Eliza Hall Institute of Medical Research, PO Royal Melbourne
            Hospital, Parkville, Victoria 3050, Australia
FEATURES             Location/Qualifiers
     source          1. .423
                     /organism="Mus musculus"
                     /db_xref="taxon:10090"
     CDS             1. .423
                     /note="pro-apoptotic BH3-containing Bcl-2 family member"
                     /codon_start=1
                     /product="BimL"
                     /protein_id="AAC40030.1"
                     /db_xref="GI:2895502"
                     /translation="MAKQPSDVSSECDREGGQLQPAERPPQLRPGAPTSLQTEPQDRS
```

```
                    PAPMSCDKSTQTPSPPCQAFNHYLSAMASIRQSQEEPEDLRPEIRIAQELRRIGDEFN
                    ETYTRRVFANDYREAEDHPQMVILQLLRFIFRLVWRRH"
BASE COUNT      113 a    116 c    109 g     85 t
ORIGIN


alignment_scores:
           Quality:  549.00             Length:    140
             Ratio:  4.991               Gaps:      1
 Percent Similarity:  78.571   Percent Identity: 78.571

alignment_block:
 US-09-508-832-2 x AF032460   ..

 Align seg 1/1  to: AF032460  from: 1  to: 423

       1 MetAlaLysGlnProSerAspValSerSerGluCysAspArgGluGlyGl 17
         ||||||||||||||||||||||||||||||||||||||||||||||||||
       1 ATGGCCAAGCAACCTTCTGATGTAAGTTCTGAGTGTGACAGAGAAGGTGG 50

      17 yGlnLeuGlnProAlaGluArgProProGlnLeuArgProGlyAlaProT 34
         ||||||||||||||||||||||||||||||||||||||||||||||||||
      51 ACAATTGCAGCCTGCTGAGAGGCCTCCCCAGCTCAGGCCTGGGGCCCCTA 100

      34 hrSerLeuGlnThrGluProGln..................................... 41
         |||||||||||||||||||||||
     101 CCTCCCTACAGACAGAACCGCAAGACAGGAGCCCGGCACCCATGAGTTGT 150

      41 ............................................................ 41

     151 GACAAGTCAACACAAACCCCAAGTCCTCCTTGCCAGGCCTTCAACCACTA 200

      42 .............AlaSerIleArgGlnSerGlnGluGluProGluAspL 54
                     ||||||||||||||||||||||||||||||||||||||
     201 TCTCAGTGCAATGGCTTCCATACGACAGTCTCAGGAGGAACCTGAAGATC 250

      54 euArgProGluIleArgIleAlaGlnGluLeuArgArgIleGlyAspGlu 70
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     251 TGCGCCCGGAGATACGGATTGCACAGGAGCTGCGGCGGATCGGAGACGAG 300

      71 PheAsnGluThrTyrThrArgArgValPheAlaAsnAspTyrArgGluAl 87
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     301 TTCAACGAAACTTACACAAGGAGGGTGTTTGCAAATGATTACCGCGAGGC 350

      87 aGluAspHisProGlnMetValIleLeuGlnLeuLeuArgPheIlePheA 104
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     351 TGAAGACCACCCTCAAATGGTTATCTTACAACTGTTACGCTTTATCTTCC 400

     104 rgLeuValTrpArgArgHis 110
         ||||||||||||||||||||
     401 GTCTGGTATGGAGAAGGCAT 420

seq_name: gb_ro:AF136927

seq_documentation_block:
LOCUS       AF136927      423 bp    mRNA            ROD       21-APR-1999
DEFINITION  Rattus norvegicus Bcl-2 related apoptotic gene product BimL (bimL)
            mRNA, complete cds.
ACCESSION   AF136927
VERSION     AF136927.1  GI:4590514
KEYWORDS    .
SOURCE      Norway rat.
  ORGANISM  Rattus norvegicus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
            Rattus.
REFERENCE   1  (bases 1 to 423)
  AUTHORS   Chen,D., Simon,R.P. and Chen,J.
  TITLE     Cloning of rat bimEL and bimL, and their differential expression in
            ischemia and normal rat brain
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 423)
  AUTHORS   Chen,D., Simon,R.P. and Chen,J.
  TITLE     Direct Submission
  JOURNAL   Submitted (24-MAR-1999) Department of Neurology, BST, S-526,
            Pittsburgh University Medical School, 3500 Terrace Street,
            Pittsburgh, PA 15213, USA
FEATURES             Location/Qualifiers
     source          1. .423
                     /organism="Rattus norvegicus"
                     /db_xref="taxon:10116"
     gene            1. .423
                     /gene="bimL"
     CDS             1. .423
                     /gene="bimL"
                     /codon_start=1
                     /product="Bcl-2 related apoptotic gene product BimL"
                     /protein_id="AAD26594.1"
                     /db_xref="GI:4590515"
                     /translation="MAKQPSDVNSECDREGGQLQPAERPPQLRPGAPTSLQTESQDRS
                     PAPMSCDKSTQTPSPPCQAFNHYLSAMASIRQSQEEPEDLRPEIRIAQELRRIGDEFN
                     ETYTRRAFANDYREAEDHPQMVILQLLRFIFRLVWRRH"
BASE COUNT      116 a    113 c    112 g     82 t
ORIGIN


alignment_scores:
           Quality:  534.00             Length:    140
             Ratio:  4.899               Gaps:      1
 Percent Similarity:  77.857   Percent Identity: 76.429

alignment_block:
 US-09-508-832-2 x AF136927   ..

 Align seg 1/1  to: AF136927  from: 1  to: 423

       1 MetAlaLysGlnProSerAspValSerSerGluCysAspArgGluGlyGl 17
         ||||||||||||||||||||||||:::||||||||||||||||||||||||
       1 ATGGCCAAGCAACCTTCTGATGTAAATTCTGAGTGTGACAGAGAAGGTGG 50

      17 yGlnLeuGlnProAlaGluArgProProGlnLeuArgProGlyAlaProT 34
         ||||||||||||||||||||||||||||||||||||||||||||||||||
      51 ACAATTGCAGCCTGCTGAGAGGCCTCCCCAGCTCAGGCCTGGGGCCCCTA 100

      34 hrSerLeuGlnThrGluProGln..................................... 41
         ||||||||||||||||||   |||
     101 CCTCCCTACAGACAGAATCGCAAGACAGGAGCCCGGCACCCATGAGTTGT 150

      41 ............................................................ 41

     151 GACAAGTCAACACAAACCCCAAGTCCTCCTTGCCAGGCCTTCAACCATTA 200

      42 .............AlaSerIleArgGlnSerGlnGluGluProGluAspL 54
                     ||||||||||||||||||||||||||||||||||||||
     201 TCTTAGCGCAATGGCTTCCATAAGGCAGTCTCAGGAGGAACCTGAAGATC 250

      54 euArgProGluIleArgIleAlaGlnGluLeuArgArgIleGlyAspGlu 70
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     251 TGCGCCCAGAGATACGGATCGCACAGGAGCTGCGGAGGATCGGAGACGAG 300

      71 PheAsnGluThrTyrThrArgArgValPheAlaAsnAspTyrArgGluAl 87
         ||||||||||||||||||||||||:::||||||||||||||||||||||||
     301 TTCAATGAGACTTACACGAGGAGGGCGTTTGCAAACGATTACCGAGAGGC 350

      87 aGluAspHisProGlnMetValIleLeuGlnLeuLeuArgPheIlePheA 104
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     351 GGAAGACCACCCGCAAATGGTTATCTTACAACTGTTACGATTCATCTTCC 400

     104 rgLeuValTrpArgArgHis 110
         ||||||||||||||||||||
     401 GTCTGGTATGGAGAAGGCAT 420

seq_name: gb_pat:AX031283

seq_documentation_block:
LOCUS       AX031283      590 bp    DNA             PAT       20-SEP-2000
```

```
DEFINITION  Sequence 5 from Patent WO9914321.
ACCESSION   AX031283
VERSION     AX031283.1  GI:10278614
KEYWORDS    .
SOURCE      unidentified.
  ORGANISM  unidentified
            unclassified.
REFERENCE   1  (bases 1 to 590)
  AUTHORS   O'Reilly,L., Puthalakath,H., Adams,J., O'Connor,L., Cory,S.,
            Huang,D.C. and Strasser,A.
  TITLE     Novel therapeutic molecules
  JOURNAL   Patent: WO 9914321-A 5 25-MAR-1999;
            INST MEDICAL W & E HALL (AU) ; PUTHALAKATH HAMSA (AU) ; REILLY
            LORRAINE O (AU) ; ADAMS JERRY (AU) ; CONNOR LIAM O (AU) ; CORY
            SUZANNE (AU) ; HUANG DAVID C S (AU) ; STRASSER ANDREAS (AU)
FEATURES             Location/Qualifiers
     source          1. .590
                     /organism="unidentified"
                     /db_xref="taxon:32644"
     CDS             1. .>587
                     /note="unnamed protein product"
                     /codon_start=1
                     /protein_id="CAC09655.1"
                     /db_xref="GI:10278615"
                     /translation="MAKQPSDVSSECDREGGQLQPAERPPQLRPGAPTSLQTEPQGNP
                     DGEGDRCPHGSPQGPLAPPASPGPFATRSPLFIFVRRSSLLSRSSSGYFSFDTDRSPA
                     PMSCDKSTQTPSPPCQAFNHYLSAMASIRQSQEEPEDLRPEIRIAQELRRIGDEFNET
                     YTRRVFANDYREAEDHPQMVILQLLRFIFRLVWRR"
BASE COUNT      137 a    178 c    150 g    125 t
ORIGIN


alignment_scores:
          Quality:  521.00            Length:    196
            Ratio:  4.736               Gaps:      1
 Percent Similarity:  56.122   Percent Identity: 56.122

alignment_block:
 US-09-508-832-2 x AX031283   ..

 Align seg 1/1  to: AX031283  from: 1  to: 590

       1 MetAlaLysGlnProSerAspValSerSerGluCysAspArgGluGlyGl 17
         |||||||||||||||||||||||||||||||||||||||||||||||||||
       1 ATGGCCAAGCAACCTTCTGATGTAAGTTCTGAGTGTGACAGAGAAGGTGG 50

      17 yGlnLeuGlnProAlaGluArgProProGlnLeuArgProGlyAlaProT 34
         ||||||||||||||||||||||||||||||||||||||||||||||||||
      51 ACAATTGCAGCCTGCTGAGAGGCCTCCCCAGCTCAGGCCTGGGGCCCCTA 100

      34 hrSerLeuGlnThrGluProGln..........................  41
         |||||||||||||||||||||||
     101 CCTCCCTACAGACAGAACCGCAAGGTAATCCCGACGGCGAAGGGGACCGC 150

      41 ...................................................  41

     151 TGCCCCCACGGCAGCCCTCAGGGCCCGCTGGCCCCACCGGCCAGCCCTGG 200

      41 ...................................................  41

     201 CCCTTTTGCTACCAGATCCCCACTTTTCATCTTTGTGAGAAGATCTTCTC 250

      41 ...................................................  41

     251 TGCTGTCCCGGTCCTCCAGTGGGTATTTCTCTTTTGACACAGACAGGAGC 300

      41 ...................................................  41

     301 CCGGCACCCATGAGTTGTGACAAGTCAACACAAACCCCAAGTCCTCCTTG 350

      42 ............................AlaSerIleArgGlnSerG  48
                                     ||||||||||||||||||||
     351 CCAGGCCTTCAACCACTATCTCAGTGCAATGGCTTCCATACGACAGTCTC 400

      48 lnGluGluProGluAspLeuArgProGluIleArgIleAlaGlnGluLeu 64
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     401 AGGAGGAACCTGAAGATCTGCGCCCGGAGATACGGATTGCACAGGAGCTG 450

      65 ArgArgIleGlyAspGluPheAsnGluThrTyrThrArgArgValPheAl 81
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     451 CGGCGGATCGGAGACGAGTTCAACGAAACTTACACAAGGAGGGTGTTTGC 500

      81 aAsnAspTyrArgGluAlaGluAspHisProGlnMetValIleLeuGlnL 98
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     501 AAATGATTACCGCGAGGCTGAAGACCACCCTCAAATGGTTATCTTACAAC 550

      98 euLeuArgPheIlePheArgLeuValTrpArgArgHis 110
         |||||||||||||||||||||||||||||||||||
     551 TGTTACGCTTTATCTTCCGTCTGGTATGGAGAAGGCAT 588

seq_name: gb_un:AX031309

seq_documentation_block:
LOCUS       AX031309      590 bp    DNA             UNA       20-SEP-2000
DEFINITION  Sequence 5 from Patent WO9914321.
ACCESSION   AX031309
VERSION     AX031309.1  GI:10278637
KEYWORDS    .
SOURCE      unidentified.
  ORGANISM  unidentified
            unclassified.
REFERENCE   1  (bases 1 to 590)
  AUTHORS   O'Reilly,L., Puthalakath,H., Adams,J., O'Connor,L., Cory,S.,
            Huang,D.C. and Strasser,A.
  TITLE     Novel therapeutic molecules
  JOURNAL   Patent: WO 9914321-A 25-MAR-1999;
            INST MEDICAL W & E HALL (AU) ; PUTHALAKATH HAMSA (AU) ; REILLY
            LORRAINE O (AU) ; ADAMS JERRY (AU) ; CONNOR LIAM O (AU) ; CORY
            SUZANNE (AU) ; HUANG DAVID C S (AU) ; STRASSER ANDREAS (AU)
FEATURES             Location/Qualifiers
     source          1. .590
                     /organism="unidentified"
                     /db_xref="taxon:32644"
     CDS             1. .>587
                     /note="unnamed protein product"
                     /codon_start=1
                     /protein_id="CAC09660.1"
                     /db_xref="GI:10278638"
                     /translation="MAKQPSDVSSECDREGGQLQPAERPPQLRPGAPTSLQTEPQGNP
                     DGEGDRCPHGSPQGPLAPPASPGPFATRSPLFIFVRRSSLLSRSSSGYFSFDTDRSPA
                     PMSCDKSTQTPSPPCQAFNHYLSAMASIRQSQEEPEDLRPEIRIAQELRRIGDEFNET
                     YTRRVFANDYREAEDHPQMVILQLLRFIFRLVWRR"
BASE COUNT      137 a    178 c    150 g    125 t
ORIGIN


alignment_scores:
          Quality:  521.00            Length:    196
            Ratio:  4.736               Gaps:      1
 Percent Similarity:  56.122   Percent Identity: 56.122

alignment_block:
 US-09-508-832-2 x AX031309   ..

 Align seg 1/1  to: AX031309  from: 1  to: 590

       1 MetAlaLysGlnProSerAspValSerSerGluCysAspArgGluGlyGl 17
         |||||||||||||||||||||||||||||||||||||||||||||||||||
       1 ATGGCCAAGCAACCTTCTGATGTAAGTTCTGAGTGTGACAGAGAAGGTGG 50

      17 yGlnLeuGlnProAlaGluArgProProGlnLeuArgProGlyAlaProT 34
         ||||||||||||||||||||||||||||||||||||||||||||||||||
      51 ACAATTGCAGCCTGCTGAGAGGCCTCCCCAGCTCAGGCCTGGGGCCCCTA 100

      34 hrSerLeuGlnThrGluProGln..........................  41
         |||||||||||||||||||||||
```

```
     101 CCTCCCTACAGACAGAACCGCAAGGTAATCCCGACGGCGAAGGGGACCGC 150

      41 ..................................................  41

     151 TGCCCCCACGGCAGCCCTCAGGGCCCGCTGGCCCCACCGGCCAGCCCTGG 200

      41 ..................................................  41

     201 CCCTTTTGCTACCAGATCCCCACTTTTCATCTTTGTGAGAAGATCTTCTC 250

      41 ..................................................  41

     251 TGCTGTCCCGGTCCTCCAGTGGGTATTTCTCTTTTGACACAGACAGGAGC 300

      41 ..................................................  41

     301 CCGGCACCCATGAGTTGTGACAAGTCAACACAAACCCCAAGTCCTCCTTG 350

      42 ..............................AlaSerIleArgGlnSerG  48
                                       ||||||||||||||||||
     351 CCAGGCCTTCAACCACTATCTCAGTGCAATGGCTTCCATACGACAGTCTC 400

      48 lnGluGluProGluAspLeuArgProGluIleArgIleAlaGlnGluLeu  64
         ||||||||||||||||||||||||||||||||||||||||||||||||||||
     401 AGGAGGAACCTGAAGATCTGCGCCCGGAGATACGGATTGCACAGGAGCTG 450

      65 ArgArgIleGlyAspGluPheAsnGluThrTyrThrArgArgValPheAl  81
         ||||||||||||||||||||||||||||||||||||||||||||||||||||
     451 CGGCGGATCGGAGACGAGTTCAACGAAACTTACACAAGGAGGGTGTTTGC 500

      81 aAsnAspTyrArgGluAlaGluAspHisProGlnMetValIleLeuGlnL  98
         ||||||||||||||||||||||||||||||||||||||||||||||||||||
     501 AAATGATTACCGCGAGGCTGAAGACCACCCTCAAATGGTTATCTTACAAC 550

      98 euLeuArgPheIlePheArgLeuValTrpArgArgHis 110
         |||||||||||||||||||||||||||||||||||
     551 TGTTACGCTTTATCTTCCGTCTGGTATGGAGAAGGCAT 588

seq_name: gb_ro:AF032459

seq_documentation_block:
LOCUS       AF032459      591 bp    mRNA            ROD       19-FEB-1998
DEFINITION  Mus musculus BimEL mRNA, complete cds.
ACCESSION   AF032459
VERSION     AF032459.1  GI:2895499
KEYWORDS    .
SOURCE      house mouse.
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 591)
  AUTHORS   O'Connor,L., Strasser,A., O'Reilly,L.A., Hausmann,G., Adams,J.M.,
            Cory,S. and Huang,D.C.
  TITLE     Bim: a novel member of the Bcl-2 family that promotes apoptosis
  JOURNAL   EMBO J. 17 (2), 384-395 (1998)
  MEDLINE   98094360
   PUBMED   9430630
REFERENCE   2  (bases 1 to 591)
  AUTHORS   O'Connor,L., Strasser,A., O'Reilly,L.A., Hausmann,G., Adams,J.M.,
            Cory,S. and Huang,D.C.S.
  TITLE     Direct Submission
  JOURNAL   Submitted (03-NOV-1997) Molecular Genetics of Cancer, The Walter &
            Eliza Hall Institute of Medical Research, PO Royal Melbourne
            Hospital, Parkville, Victoria 3050, Australia
FEATURES             Location/Qualifiers
     source          1. .591
                     /organism="Mus musculus"
                     /db_xref="taxon:10090"
     CDS             1. .591
                     /note="pro-apoptotic BH3-containing Bcl-2 family member"
                     /codon_start=1
                     /product="BimEL"
                     /protein_id="AAC40029.1"
                     /db_xref="GI:2895500"
                     /translation="MAKQPSDVSSECDREGGQLQPAERPPQLRPGAPTSLQTEPQGNP
                     DGEGDRCPHGSPQGPLAPPASPGPFATRSPLFIFVRRSSLLSRSSSGYFSFDTDRSPA
                     PMSCDKSTQTPSPPCQAFNHYLSAMASIRQSQEEPEDLRPEIRIAQELRRIGDEFNET
                     YTRRVFANDYREAEDHPQMVILQLLRFIFRLVWRRH"
BASE COUNT      138 a    178 c    150 g    125 t
ORIGIN


alignment_scores:
           Quality:  521.00            Length:    196
             Ratio:  4.736               Gaps:      1
 Percent Similarity:  56.122   Percent Identity: 56.122

alignment_block:
 US-09-508-832-2 x AF032459    ..

 Align seg 1/1  to: AF032459  from: 1  to: 591

       1 MetAlaLysGlnProSerAspValSerSerGluCysAspArgGluGlyGl 17
         ||||||||||||||||||||||||||||||||||||||||||||||||||||
       1 ATGGCCAAGCAACCTTCTGATGTAAGTTCTGAGTGTGACAGAGAAGGTGG 50

      17 yGlnLeuGlnProAlaGluArgProProGlnLeuArgProGlyAlaProT 34
         ||||||||||||||||||||||||||||||||||||||||||||||||||||
      51 ACAATTGCAGCCTGCTGAGAGGCCTCCCCAGCTCAGGCCTGGGGCCCCTA 100

      34 hrSerLeuGlnThrGluProGln..................................  41
         |||||||||||||||||||||||
     101 CCTCCCTACAGACAGAACCGCAAGGTAATCCCGACGGCGAAGGGGACCGC 150

      41 ..................................................  41

     151 TGCCCCCACGGCAGCCCTCAGGGCCCGCTGGCCCCACCGGCCAGCCCTGG 200

      41 ..................................................  41

     201 CCCTTTTGCTACCAGATCCCCACTTTTCATCTTTGTGAGAAGATCTTCTC 250

      41 ..................................................  41

     251 TGCTGTCCCGGTCCTCCAGTGGGTATTTCTCTTTTGACACAGACAGGAGC 300

      41 ..................................................  41

     301 CCGGCACCCATGAGTTGTGACAAGTCAACACAAACCCCAAGTCCTCCTTG 350

      42 ..............................AlaSerIleArgGlnSerG  48
                                       ||||||||||||||||||
     351 CCAGGCCTTCAACCACTATCTCAGTGCAATGGCTTCCATACGACAGTCTC 400

      48 lnGluGluProGluAspLeuArgProGluIleArgIleAlaGlnGluLeu  64
         ||||||||||||||||||||||||||||||||||||||||||||||||||||
     401 AGGAGGAACCTGAAGATCTGCGCCCGGAGATACGGATTGCACAGGAGCTG 450

      65 ArgArgIleGlyAspGluPheAsnGluThrTyrThrArgArgValPheAl  81
         ||||||||||||||||||||||||||||||||||||||||||||||||||||
     451 CGGCGGATCGGAGACGAGTTCAACGAAACTTACACAAGGAGGGTGTTTGC 500

      81 aAsnAspTyrArgGluAlaGluAspHisProGlnMetValIleLeuGlnL  98
         ||||||||||||||||||||||||||||||||||||||||||||||||||||
     501 AAATGATTACCGCGAGGCTGAAGACCACCCTCAAATGGTTATCTTACAAC 550

      98 euLeuArgPheIlePheArgLeuValTrpArgArgHis 110
         |||||||||||||||||||||||||||||||||||
     551 TGTTACGCTTTATCTTCCGTCTGGTATGGAGAAGGCAT 588

seq_name: gb_ro:AF065433

seq_documentation_block:
LOCUS       AF065433      591 bp    mRNA            ROD       11-MAR-1999
DEFINITION  Rattus norvegicus Bcl-2 related ovarian death gene product BOD-L
            mRNA, complete cds.
```

```
ACCESSION   AF065433
VERSION     AF065433.1  GI:3228569
KEYWORDS    .
SOURCE      Norway rat.
  ORGANISM  Rattus norvegicus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
            Rattus.
REFERENCE   1  (bases 1 to 591)
  AUTHORS   Hsu,S.Y., Lin,P. and Hsueh,A.J.
  TITLE     BOD (Bcl-2-related ovarian death gene) is an ovarian BH3
            domain-containing proapoptotic Bcl-2 protein capable of
            dimerization with diverse antiapoptotic Bcl-2 members
  JOURNAL   Mol. Endocrinol. 12 (9), 1432-1440 (1998)
  MEDLINE   98400436
REFERENCE   2  (bases 1 to 591)
  AUTHORS   Hsu,S.Y. and Hsueh,A.J.W.
  TITLE     Direct Submission
  JOURNAL   Submitted (15-MAY-1998) GYN/OB, Stanford University, MSOB S385,
            Stanford, CA 94305, USA
FEATURES             Location/Qualifiers
     source          1. .591
                     /organism="Rattus norvegicus"
                     /db_xref="taxon:10116"
     CDS             1. .591
                     /codon_start=1
                     /product="Bcl-2 related ovarian death gene product BOD-L"
                     /protein_id="AAC23595.1"
                     /db_xref="GI:3228570"
                     /translation="MAKQPSDVNSECDREGGQLQPAERPPQLRPGAPTSLQTESQGNP
                     DGEGDRCPHGSPQGPLAPPASPGPFATRSPLFIFVRRSSLLSRSSSGYFSFDTDRSPA
                     PMSCDKSTQTPSPPCQAFNHYLSAMASIRQSQEEPEDLRPEIRIAQELRRIGDEFNET
                     YTRRAFANDYREAEDHPQMVILQLLRFIFRLVWRRH"
BASE COUNT      139 a    177 c    153 g    122 t
ORIGIN

alignment_scores:
            Quality:  506.00            Length:    196
              Ratio:  4.642               Gaps:      1
 Percent Similarity:  55.612   Percent Identity: 54.592

alignment_block:
 US-09-508-832-2 x AF065433   ..

 Align seg 1/1  to: AF065433  from: 1  to: 591

       1 MetAlaLysGlnProSerAspValSerSerGluCysAspArgGluGlyGl 17
         ||||||||||||||||||||||:::||||||||||||||||||||||||
       1 ATGGCCAAGCAACCTTCTGATGTAAATTCTGAGTGTGACAGAGAAGGTGG 50

      17 yGlnLeuGlnProAlaGluArgProProGlnLeuArgProGlyAlaProT 34
         ||||||||||||||||||||||||||||||||||||||||||||||||||
      51 ACAATTGCAGCCTGCTGAGAGGCCTCCCCAGCTCAGGCCTGGGGCCCCTA 100

      34 hrSerLeuGlnThrGluProGln..................... 41
         ||||||||||||||||   |||
     101 CCTCCCTACAGACAGAATCGCAAGGTAATCCCGACGGCGAAGGGGACCGC 150

      41 ............................................. 41

     151 TGCCCCCACGGCAGCCCTCAGGGCCCGCTGGCCCCACCGGCCAGCCCTGG 200

      41 ............................................. 41

     201 TCCTTTTGCTACCAGATCCCCACTTTTCATCTTTGTGAGAAGATCTTCTC 250

      41 ............................................. 41

     251 TGCTGTCCCGGTCCTCCAGTGGGTATTTCTCTTTTGACACAGACAGGAGC 300

      41 ............................................. 41

     301 CCGGCACCCATGAGTTGTGACAAGTCAACACAAACCCCAAGTCCTCCTTG 350

      42 ..............................AlaSerIleArgGlnSerG 48
                                       ||||||||||||||||||
     351 CCAGGCCTTCAACCATTATCTCAGTGCAATGGCTTCCATAAGGCAGTCTC 400

      48 lnGluGluProGluAspLeuArgProGluIleArgIleAlaGlnGluLeu 64
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     401 AGGAGGAACCTGAAGATCTGCGCCCAGAGATACGGATCGCACAGGAGCTG 450

      65 ArgArgIleGlyAspGluPheAsnGluThrTyrThrArgArgValPheAl 81
         ||||||||||||||||||||||||||||||||||||||||||:::|||||
     451 CGGCGGATCGGAGACGAGTTCAATGAGACTTACACGAGGAGGGCGTTTGC 500

      81 aAsnAspTyrArgGluAlaGluAspHisProGlnMetValIleLeuGlnL 98
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     501 AAACGATTACCGAGAGGCGGAAGACCACCCGCAAATGGTTATCTTACAAC 550

      98 euLeuArgPheIlePheArgLeuValTrpArgArgHis 110
         |||||||||||||||||||||||||||||||||||
     551 TGTTACGATTCATCTTCCGTCTGGTCTGGAGAAGGCAC 588

seq_name: gb_pat:AX031285

seq_documentation_block:
LOCUS       AX031285      416 bp    DNA             PAT       20-SEP-2000
DEFINITION  Sequence 7 from Patent WO9914321.
ACCESSION   AX031285
VERSION     AX031285.1  GI:10278616
KEYWORDS    .
SOURCE      unidentified.
  ORGANISM  unidentified
            unclassified.
REFERENCE   1  (bases 1 to 416)
  AUTHORS   O'Reilly,L., Puthalakath,H., Adams,J., O'Connor,L., Cory,S.,
            Huang,D.C. and Strasser,A.
  TITLE     Novel therapeutic molecules
  JOURNAL   Patent: WO 9914321-A 7 25-MAR-1999;
            INST MEDICAL W & E HALL (AU) ; PUTHALAKATH HAMSA (AU) ; REILLY
            LORRAINE O (AU) ; ADAMS JERRY (AU) ; CONNOR LIAM O (AU) ; CORY
            SUZANNE (AU) ; HUANG DAVID C S (AU) ; STRASSER ANDREAS (AU)
FEATURES             Location/Qualifiers
     source          1. .416
                     /organism="unidentified"
                     /db_xref="taxon:32644"
     CDS             1. .>414
                     /note="unnamed protein product"
                     /codon_start=1
                     /protein_id="CAC09656.1"
                     /db_xref="GI:10278617"
                     /translation="MAKQPSDVSSECDREGRQLQPAERPPQLRPGAPTSLQTEPQDRS
                     PAPMSCDKSTQTPSPPCQAFNHYLSAMASMRQAEPADMRPEIWIAQELRRIGDEFNAY
                     YARRVFLNNYQAAEDHPRMVILRLLRYIVRLVWRMH"
BASE COUNT      113 a    113 c    103 g     87 t
ORIGIN

alignment_scores:
            Quality:  436.00            Length:    140
              Ratio:  4.404               Gaps:      2
 Percent Similarity:  70.714   Percent Identity: 64.286

alignment_block:
 US-09-508-832-2 x AX031285   ..

 Align seg 1/1  to: AX031285  from: 1  to: 416

       1 MetAlaLysGlnProSerAspValSerSerGluCysAspArgGluGlyGl 17
         ||||||||||||||||||||||||||||||||||||||||||||||||||
       1 ATGGCAAAGCAACCTTCTGATGTAAGTTCTGAGTGTGACCGAGAAGGTAG 50

      17 yGlnLeuGlnProAlaGluArgProProGlnLeuArgProGlyAlaProT 34
         ||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
      51 ACAATTGCAGCCTGCGGAGAGGCCTCCCCAGCTCAGACCTGGGGCCCCTA 100

      34 hrSerLeuGlnThrGluProGln..........................  41
         ||||||||||||||||||||||
     101 CCTCCCTACAGACAGAGCCACAAGACAGGAGCCCAGCACCCATGAGTTGT 150

      41 .................................................  41

     151 GACAAATCAACACAAACCCCAAGTCCTCCTTGCCAGGCCTTCAACCACTA 200

      42 .............AlaSerIleArgGlnSerGlnGluGluProGluAspL  54
                     ||||||:::||||||:::      ||||||   |||:
     201 TCTCAGTGCAATGGCTTCCATGAGGCAGGCT......GAACCTGCAGATA 244

      54 euArgProGluIleArgIleAlaGlnGluLeuArgArgIleGlyAspGlu  70
         ::|||||||||||||   ||||||||||||||||||||||||||||||||
     245 TGCGCCCAGAGATATGGATCGCCCAAGAGTTGCGGCGTATCGGAGACGAG 294

      71 PheAsnGluThrTyrThrArgArgValPheAlaAsnAspTyrArgGluAl  87
         ||||||      |||:::||||||||||||   |||:::|||:::   ||
     295 TTTAACGCTTACTATGCAAGGAGGGTATTTTTGAATAATTACCAAGCAGC 344

      87 aGluAspHisProGlnMetValIleLeuGlnLeuLeuArgPheIlePheA 104
         ||||||||||||||:::||||||||||||:::||||||||||:::|||   |
     345 CGAAGACCACCCACGAATGGTTATCTTACGACTGTTACGTTACATTGTCC 394

     104 rgLeuValTrpArgArgHis 110
         ||||||||||||||   |||
     395 GCCTGGTGTGGAGAATGCAT 414

seq_name: gb_un:AX031311

seq_documentation_block:
LOCUS       AX031311      416 bp    DNA             UNA       20-SEP-2000
DEFINITION  Sequence 7 from Patent WO9914321.
ACCESSION   AX031311
VERSION     AX031311.1  GI:10278639
KEYWORDS    .
SOURCE      unidentified.
  ORGANISM  unidentified
            unclassified.
REFERENCE   1  (bases 1 to 416)
  AUTHORS   O'Reilly,L., Puthalakath,H., Adams,J., O'Connor,L., Cory,S.,
            Huang,D.C. and Strasser,A.
  TITLE     Novel therapeutic molecules
  JOURNAL   Patent: WO 9914321-A 25-MAR-1999;
            INST MEDICAL W & E HALL (AU) ; PUTHALAKATH HAMSA (AU) ; REILLY
            LORRAINE O (AU) ; ADAMS JERRY (AU) ; CONNOR LIAM O (AU) ; CORY
            SUZANNE (AU) ; HUANG DAVID C S (AU) ; STRASSER ANDREAS (AU)
FEATURES             Location/Qualifiers
     source          1. .416
                     /organism="unidentified"
                     /db_xref="taxon:32644"
     CDS             1. .>414
                     /note="unnamed protein product"
                     /codon_start=1
                     /protein_id="CAC09661.1"
                     /db_xref="GI:10278640"
                     /translation="MAKQPSDVSSECDREGRQLQPAERPPQLRPGAPTSLQTEPQDRS
                     PAPMSCDKSTQTPSPPCQAFNHYLSAMASMRQAEPADMRPEIWIAQELRRIGDEFNAY
                     YARRVFLNNYQAAEDHPRMVILRLLRYIVRLVWRMH"
BASE COUNT      113 a    113 c    103 g     87 t
ORIGIN


alignment_scores:
            Quality:  436.00            Length:    140
              Ratio:  4.404               Gaps:      2
 Percent Similarity:  70.714   Percent Identity: 64.286

alignment_block:
 US-09-508-832-2 x AX031311   ..
```

```
 Align seg 1/1  to: AX031311  from: 1  to: 416

       1 MetAlaLysGlnProSerAspValSerSerGluCysAspArgGluGlyGl 17
         ||||||||||||||||||||||||||||||||||||||||||||||||||||
       1 ATGGCAAAGCAACCTTCTGATGTAAGTTCTGAGTGTGACCGAGAAGGTAG 50

      17 yGlnLeuGlnProAlaGluArgProProGlnLeuArgProGlyAlaProT 34
         ||||||||||||||||||||||||||||||||||||||||||||||||||||
      51 ACAATTGCAGCCTGCGGAGAGGCCTCCCCAGCTCAGACCTGGGGCCCCTA 100

      34 hrSerLeuGlnThrGluProGln..........................  41
         ||||||||||||||||||||||
     101 CCTCCCTACAGACAGAGCCACAAGACAGGAGCCCAGCACCCATGAGTTGT 150

      41 .................................................  41

     151 GACAAATCAACACAAACCCCAAGTCCTCCTTGCCAGGCCTTCAACCACTA 200

      42 .............AlaSerIleArgGlnSerGlnGluGluProGluAspL  54
                     ||||||:::||||||:::      ||||||   |||:
     201 TCTCAGTGCAATGGCTTCCATGAGGCAGGCT......GAACCTGCAGATA 244

      54 euArgProGluIleArgIleAlaGlnGluLeuArgArgIleGlyAspGlu  70
         ::|||||||||||||   ||||||||||||||||||||||||||||||||
     245 TGCGCCCAGAGATATGGATCGCCCAAGAGTTGCGGCGTATCGGAGACGAG 294

      71 PheAsnGluThrTyrThrArgArgValPheAlaAsnAspTyrArgGluAl  87
         ||||||      |||:::||||||||||||   |||:::|||:::   ||
     295 TTTAACGCTTACTATGCAAGGAGGGTATTTTTGAATAATTACCAAGCAGC 344

      87 aGluAspHisProGlnMetValIleLeuGlnLeuLeuArgPheIlePheA 104
         ||||||||||||||:::||||||||||||:::||||||||||:::|||   |
     345 CGAAGACCACCCACGAATGGTTATCTTACGACTGTTACGTTACATTGTCC 394

     104 rgLeuValTrpArgArgHis 110
         ||||||||||||||   |||
     395 GCCTGGTGTGGAGAATGCAT 414

seq_name: gb_pr:AF032458

seq_documentation_block:
LOCUS       AF032458      417 bp    mRNA            PRI       19-FEB-1998
DEFINITION  Homo sapiens BimL mRNA, complete cds.
ACCESSION   AF032458
VERSION     AF032458.1  GI:2895497
KEYWORDS    .
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 417)
  AUTHORS   O'Connor,L., Strasser,A., O'Reilly,L.A., Hausmann,G., Adams,J.M.,
            Cory,S. and Huang,D.C.
  TITLE     Bim: a novel member of the Bcl-2 family that promotes apoptosis
  JOURNAL   EMBO J. 17 (2), 384-395 (1998)
  MEDLINE   98094360
   PUBMED   9430630
REFERENCE   2  (bases 1 to 417)
  AUTHORS   O'Connor,L., Strasser,A., O'Reilly,L.A., Hausmann,G., Adams,J.M.,
            Cory,S. and Huang,D.C.S.
  TITLE     Direct Submission
  JOURNAL   Submitted (03-NOV-1997) Molecular Genetics of Cancer, The Walter &
            Eliza Hall Institute of Medical Research, PO Royal Melbourne
            Hospital, Parkville, Victoria 3050, Australia
FEATURES             Location/Qualifiers
     source          1. .417
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
     CDS             1. .417
                     /note="pro-apoptotic BH3-containing Bcl-2 family member"
                     /codon_start=1
                     /product="BimL"
                     /protein_id="AAC39594.1"
```

```
                     /db_xref="GI:2895498"
                     /translation="MAKQPSDVSSECDREGRQLQPAERPPQLRPGAPTSLQTEPQDRS
                     PAPMSCDKSTQTPSPPCQAFNHYLSAMASMRQAEPADMRPEIWIAQELRRIGDEFNAY
                     YARRVFLNNYQAAEDHPRMVILRLLRYIVRLVWRMH"
BASE COUNT      114 a    113 c    103 g     87 t
ORIGIN


alignment_scores:
            Quality:  436.00             Length:    140
              Ratio:  4.404                Gaps:      2
 Percent Similarity:  70.714   Percent Identity: 64.286

alignment_block:
 US-09-508-832-2 x AF032458   ..

 Align seg 1/1  to: AF032458  from: 1  to: 417

       1 MetAlaLysGlnProSerAspValSerSerGluCysAspArgGluGlyGl 17
         ||||||||||||||||||||||||||||||||||||||||||||||||||
       1 ATGGCAAAGCAACCTTCTGATGTAAGTTCTGAGTGTGACCGAGAAGGTAG 50

      17 yGlnLeuGlnProAlaGluArgProProGlnLeuArgProGlyAlaProT 34
         ||||||||||||||||||||||||||||||||||||||||||||||||||
      51 ACAATTGCAGCCTGCGGAGAGGCCTCCCCAGCTCAGACCTGGGGCCCCTA 100

      34 hrSerLeuGlnThrGluProGln..................................... 41
         ||||||||||||||||||||||
     101 CCTCCCTACAGACAGAGCCACAAGACAGGAGCCCAGCACCCATGAGTTGT 150

      41 ............................................................ 41

     151 GACAAATCAACACAAACCCCAAGTCCTCCTTGCCAGGCCTTCAACCACTA 200

      Q2 ............AlaSerIleArgGlnSerGlnGluGluProGluAspL 54
                     ||||||:::|||||||:::      ||||||   |||:
     201 TCTCAGTGCAATGGCTTCCATGAGGCAGGCT......GAACCTGCAGATA 244

      54 euArgProGluIleArgIleAlaGlnGluLeuArgArgIleGlyAspGlu 70
         ::|||||||||||   ||||||||||||||||||||||||||||||||||
     245 TGCGCCCAGAGATATGGATCGCCCAAGAGTTGCGGCGTATCGGAGACGAG 294

      71 PheAsnGluThrTyrThrArgArgValPheAlaAsnAspTyrArgGluAl 87
         ||||||      |||:::|||||||||||||   |||:::|||:::   ||
     295 TTTAACGCTTACTATGCAAGGAGGGTATTTTTGAATAATTACCAAGCAGC 344

      87 aGluAspHisProGlnMetValIleLeuGlnLeuLeuArgPheIlePheA 104
         |||||||||||||:::|||||||||||||:::||||||||||:::|||   |
     345 CGAAGACCACCCACGAATGGTTATCTTACGACTGTTACGTTACATTGTCC 394

     104 rgLeuValTrpArgArgHis 110
         ||||||||||||||   |||
     395 GCCTGGTGTGGAGAATGCAT 414
```

Walz M;

6/32.

2

ubulin from Acremonium chrysogenum - and relates DNA, nd transformed cells, for co-transformations of a wide A. chrysogenum strains with foreign genes

laim 4; Page 9-10; 31pp; German.

Vectors based upon the beta tubulin coding sequence may be used to transform Acremonium chrysogenum, optionally in combination with other vectors to introduce a required foreign gene such as the gene encoding glutaryl acylase, an enzyme involved in cephalosporin biosynthesis. The coding sequence may be used to transform a wide variety of Acremonium chrysogenum strains (wild type and mutants). Unlike known systems, it is not recipient-strain limited.

SQ Sequence 2206 BP; 449 A; 740 C; 541 G; 476 T; 0 other;

```
alignment_scores:
            Quality:   84.50              Length:     95
              Ratio:  1.536                Gaps:      6
 Percent Similarity:  57.895   Percent Identity: 30.526

alignment_block:
 US-09-508-832-2 x AAQ70754    ..

 Align seg 1/1  to: AAQ70754  from: 1  to: 2206

      15 GluGlyGlyGlnLeuGlnProAlaGluArgProProGln..... 27
         ?|||||||||:::          |||:::|||||   |||
    1680 GAAGGAGGTCGAGGACCAGATGCGCAACGTCCAGAGCAAGAACTCGTCCT 1729

      28 .LeuArgProGlyAlaProThrSerLeuGlnThrGluProGlnAlaSerI 44
          ||||||         |||         |||   :::|||:::   |||:
    1730 ACTTCGTCGAGTGGATCCC.....CAACAACATCCAGACCGCTCTC 1770

      44 leArgGlnSerGlnGluGluProGluAsp.....LeuArgProGlu 57
         ::|||:::|||:::      |||:::|||        |||:::   :::
    1771 TGCGCCATTCCTCCCCGTGGCCTCAAGATGTCCTCCACCTTCATCGGCAA 1820

      58 IleArgIle.AlaGlnGluLeu...ArgArgIleGlyAspGluPheAsnG 73
         :::::::::    |||||||||   :::|||:::|||:::::|||:::
    1821 CTCCACCTCCATCCAGGAGCTGTTCAAGCGTGTCGGTGAGCAGTTCACTG 1870

      73 luThrTyrThrArgArgValPheAlaAsn..................... 82
            :::   |||:::::::|||   :::
    1871 CCATGTTCCGTCGCAAGGCTTTCCTGCATTGGTACACTGGTGAGGGCATG 1920

      83 .........AspTyrArgGluAlaGluAspHis 90
                 ::::::   |||||||||:::::
    1921 GACGAGATGGAGTTTACCGAGGCCGAGTCCAAC 1953

seq_name: /SIDS2/gcgdata/geneseq/geneseqn/NA1993.DAT:AAQ48230

seq_documentation_block:
ID   AAQ48230 standard; DNA; 3445 BP.
XX
AC   AAQ48230;
XX
DT   22-FEB-1994  (first entry)
XX
DE   Acremonium chrysogenum beta-tubulin gene.
XX
KW   Beta tubulin; mutant; chemical resistance; selective marker;
KW   cephalosporin; antibiotic production; ds.
XX
OS   Acremonium chrysogenum.
XX
FH   Key             Location/Qualifiers
```

```
FT   exon            1..1298
FT                   /*tag= a
FT                   /number= 1
FT                   /note= "ATG initiation codon is located at
FT                        nucleotides 1287..1289"
FT   intron          1299..1460
FT                   /*tag= b
FT                   /number= 1
FT   exon            1461..1484
FT                   /*tag= c
FT                   /number= 2
FT   intron          1485..1551
FT                   /*tag= d
FT                   /number= 2
FT   exon            1552..1674
FT                   /*tag= e
FT                   /number= 3
FT   intron          1675..1748
FT                   /*tag= f
FT                   /number= 3
FT   exon            1749..2539
FT                   /*tag= g
FT                   /number= 4
FT   intron          2540..2602
FT                   /*tag= h
FT                   /number= 4
FT   exon            2603..3445
FT                   /*tag= i
FT                   /number= 5
FT                   /note= "TAA termination codon is located at
FT                        nucleotides 2994..2996"
XX
PN   JP05192157-A.
XX
PD   03-AUG-1993.
XX
PF   26-MAY-1992;   92JP-0133384.
XX
PR   27-MAY-1991;   91JP-0121276.
XX
PA   (TAKE ) TAKEDA CHEM IND LTD.
XX
DR   WPI; 1993-277472/35.
DR   P-PSDB; AAR40226.
XX
PT   DNA fragment contg. DNA coding mutant beta-tubulin - originates
PT   from Acremonium chrysogenum, used as selective marker for
PT   transformation of A.chrysogenum
XX
PS   Example 5; Fig 4-6; 16pp; Japanese.
XX
CC   The wild-type coding sequence for beta-tubulin was isolated from
CC   Acremonium chrysogenum ATCC 11550 and sequenced (AAQ48230). Primers
CC   CTU-3 and CTU-6 were used to introduce mutations at codon 100 (Asn
CC   to Ile) and 167 (Phe to Tyr), respectively. Expression of the
CC   mutant proteins encoded by these sequences confers chemical
CC   resistance (e.g. to carbendazim and to ansamitocin) on transformed
CC   microorganisms. See AAQ55405 and AAQ55406 for mutated sequences.
XX
SQ   Sequence 3445 BP; 723 A; 1061 C; 892 G; 769 T; 0 other;

alignment_scores:
            Quality:   84.50              Length:     95
              Ratio:  1.536                Gaps:      6
 Percent Similarity:  57.895   Percent Identity: 30.526

alignment_block:
 US-09-508-832-2 x AAQ48230    ..

 Align seg 1/1  to: AAQ48230  from: 1  to: 3445

      15 GluGlyGlyGlnLeuGlnProAlaGluArgProProGln..... 27
```

File 3

```
    1194 AAGTCGAGCGTGTGCGACATCCCGCCCAGGGGCCTGAAGATGGCGGGGAC 1243

      54 LeuArgProGluIleArgIle.AlaGlnGluLeu...ArgArgIleGlyA 69
         ||||||   |||:::::::::    |||||||::   ||||||:::::::
    1244 CTTCGTGGGGAACTCCACCTCCATCCAGGAGATGTTCCGCAGGGTGAGCG 1293

      69 spGluPheAsnGluThrTyrThrArgArgValPheAlaAsn..... 82
         :::::|||:::      :::   |||:::::::|||   :::
    1294 AGCAGTTCACCGCCATGTTCAGGCGCAAGGCCTTCCTGCACTGGTACACC 1343

      83 .....AspTyrArgGluAlaGluAspHisProGl 92
                ::::::   |||||||||:::::::   ::
    1344 GGCGAGGGGATGGATGAGATGGAGTTCACTGAGGCCGAGAGCAACATGAA 1393

      92 nMetValIleLeuGlnLeuLeuArgPhe 101
         :   ::::::   :::      ::::::
    1394 CGATCTGGTCGCCGAGTACCAGCAGTAC 1421

seq_name: /SIDS2/gcgdata/geneseq/geneseqn/NA1994.DAT:AAQ70754

seq_documentation_block:
ID   AAQ70754 standard; DNA; 2206 BP.
XX
AC   AAQ70754;
XX
DT   22-MAR-1995  (first entry)
XX
DE   Beta tubulin gene of Acremonium chrysogenum.
XX
KW   Beta tubulin; Acremonium chrysogenum; transformation;
KW   co-transformation; biosynthesis; cephalosporins; ss.
XX
OS   Acremonium chrysogenum.
XX
FH   Key             Location/Qualifiers
FT   exon            346..357
FT                   /*tag= a
FT                   /label= Exon 1.
FT   intron          358..519
FT                   /*tag= b
FT                   /label= Intron 1.
FT   exon            520..543
FT                   /*tag= c
FT                   /label= Exon 2.
FT   intron          544..610
FT                   /*tag= d
FT                   /label= Intron 2.
FT   exon            611..733
FT                   /*tag= e
FT                   /label= Exon 3.
FT   intron          734..807
FT                   /*tag= f
FT                   /label= Intron 3.
FT   exon            808..1598
FT                   /*tag= g
FT                   /label= Exon 4.
FT   intron          1599..1661
FT                   /*tag= h
FT                   /label= Intron 4.
FT   exon            1662..2055
FT                   /*tag= i
FT                   /label= Exon 5.
XX
PN   EP610842-A.
XX
PD   17-AUG-1994.
XX
PF   07-FEB-1994;   94EP-0101794.
XX
PR   12-FEB-1993;   93DE-4304312.
XX
PA   (FARH ) HOECHST AG.
XX
```

```
OM of: US-09-508-832-2 to: N_Geneseq_1101:*   out_format : pfs

Date: Dec 11, 2001  1:49 AM

About: Results were produced by the GenCore software, version 4.5,
       Copyright (c) 1993-2000 Compugen Ltd.

Command line parameters:
 -MODEL=frame+_p2n.model -DEV=xlp
 -Q=/cgn2_1/USPTO_spool/US09508832/runat_10122001_110349_29569/app_query.fasta_1.620
 -DB=N_Geneseq_1101 -QFMT=fastap -SUFFIX=rng -GAPOP=12.000
 -GAPEXT=4.000 -MINMATCH=0.100 -LOOPCL=0.000 -LOOPEXT=0.000
 -QGAPOP=4.500 -QGAPEXT=0.050 -XGAPOP=10.000 -XGAPEXT=0.500
 -FGAPOP=6.000 -FGAPEXT=7.000 -YGAPOP=10.000 -YGAPEXT=0.500
 -DELOP=6.000 -DELEXT=7.000 -START=1 -MATRIX=blosum62
 -TRANS=human40.cdi -LIST=45 -DOCALIGN=200 -THR_SCORE=pct
 -THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=pfs
 -NORM=ext -MINLEN=0 -MAXLEN=2000000000
 -USER=US09508832_@CGN1_1_327 -NCPU=6 -ICPU=3 -LONGLOG -NO_XLPXY
 -WAIT -THREADS=1

Search information block:
Query: US-09-508-832-2
Query length: 110
Database: N_Geneseq_1101:*
Database sequences: 930621
Database length: 428662619
Search time (sec): 249.590000



score_list:
Sequence              Strd Orig    ZScore    EScore Len  ! Documentation  ..
/SIDS2/gcgdata/geneseq/geneseqn/NA1999.DAT:AAX24993 +  574.00 1102.40  2.5e-53   332  !
/SIDS2/gcgdata/geneseq/geneseqn/NA1999.DAT:AAX24994 +  549.00 1051.96  1.6e-50   422  !
/SIDS2/gcgdata/geneseq/geneseqn/NA1999.DAT:AAX24995 +  521.00  994.85  2.5e-47   590  !
/SIDS2/gcgdata/geneseq/geneseqn/NA1999.DAT:AAX24996 +  436.00  834.40  2.1e-38   416  !
/SIDS2/gcgdata/geneseq/geneseqn/NA1999.DAT:AAX24997 +  406.00  773.20  5.4e-35   596  !
/SIDS2/gcgdata/geneseq/geneseqn/NA2000.DAT:AAC44501 +   87.50  149.42  3.02     1747  !
/SIDS2/gcgdata/geneseq/geneseqn/NA1994.DAT:AAQ70754 +   84.50  141.43  8.41     2206  !
/SIDS2/gcgdata/geneseq/geneseqn/NA1993.DAT:AAQ48230 +   84.50  137.21  14.46    3445  !
/SIDS2/gcgdata/geneseq/geneseqn/NA1993.DAT:AAQ55405 +   84.50  137.21  14.46    3445  !
/SIDS2/gcgdata/geneseq/geneseqn/NA1993.DAT:AAQ55406 +   84.50  137.21  14.46    3445  !
/SIDS2/gcgdata/geneseq/geneseqn/NA2000.DAT:AAA63350 -   83.00  117.12  190.26   21185
/SIDS2/gcgdata/geneseq/geneseqn/NA2000.DAT:AAA63348 -   83.00  106.77  717.21   63164
/SIDS2/gcgdata/geneseq/geneseqn/NA2000.DAT:AAZ90014 +   79.50  137.90  13.23    1158  !
/SIDS2/gcgdata/geneseq/geneseqn/NA2001.DAT:AAH50546 +   79.00  131.37  30.58    2085  !
/SIDS2/gcgdata/geneseq/geneseqn/NA2001.DAT:AAH31146 +   79.00  131.37  30.58    2085  !
/SIDS2/gcgdata/geneseq/geneseqn/NA2001.DAT:AAH31243 +   79.00  131.37  30.58    2085  !
/SIDS2/gcgdata/geneseq/geneseqn/NA2001.DAT:AAD04443 +   79.00  131.37  30.58    2085  !
/SIDS2/gcgdata/geneseq/geneseqn/NA2001.DAT:AAI61371 +   79.00   83.24  1.5e+04  335913
/SIDS2/gcgdata/geneseq/geneseqn/NA2001.DAT:AAI61372 +   79.00   83.24  1.5e+04  335913
/SIDS2/gcgdata/geneseq/geneseqn/NA2000.DAT:AAC96918 +   78.00  132.75  25.63    1471  !
/SIDS2/gcgdata/geneseq/geneseqn/NA2001.DAT:AAI24461 -   77.00  139.36  10.97    597  !
/SIDS2/gcgdata/geneseq/geneseqn/NA2001.DAT:AAI15287 -   77.00  128.04  46.85    1972  !
/SIDS2/gcgdata/geneseq/geneseqn/NA1998.DAT:AAV73325 +   76.50  131.10  31.66    1290  !
/SIDS2/gcgdata/geneseq/geneseqn/NA1999.DAT:AAX32491 +   76.50  130.28  35.18    1407  !
/SIDS2/gcgdata/geneseq/geneseqn/NA2001.DAT:AAH02918 +   76.50  130.25  35.33    1412  !
/SIDS2/gcgdata/geneseq/geneseqn/NA2001.DAT:AAH33759 +   76.50  130.15  35.79    1427  !
/SIDS2/gcgdata/geneseq/geneseqn/NA2000.DAT:AAA70056 -   76.00  143.14  6.76     327  !
/SIDS2/gcgdata/geneseq/geneseqn/NA2000.DAT:AAA77741 -   76.00  143.02  6.86     331  !
/SIDS2/gcgdata/geneseq/geneseqn/NA2001.DAT:AAI28479 -   76.00  143.02  6.86     331  !
/SIDS2/gcgdata/geneseq/geneseqn/NA2001.DAT:AAH83514 +   76.00  140.81  9.11     418  !
/SIDS2/gcgdata/geneseq/geneseqn/NA1997.DAT:AAT91561 +   76.00  133.10  24.51    944  !
/SIDS2/gcgdata/geneseq/geneseqn/NA1988.DAT:AAN80309 +   76.00  124.70  71.96    2291  !
/SIDS2/gcgdata/geneseq/geneseqn/NA1993.DAT:AAQ38230 +   75.50  129.14  40.72    1295  !
/SIDS2/gcgdata/geneseq/geneseqn/NA1998.DAT:AAV42919 +   75.50  128.65  43.33    1363  !
/SIDS2/gcgdata/geneseq/geneseqn/NA1998.DAT:AAV42918 +   75.50  128.61  43.56    1369  !
/SIDS2/gcgdata/geneseq/geneseqn/NA2000.DAT:AAZ98405 +   75.50  127.59  49.67    1525  !
/SIDS2/gcgdata/geneseq/geneseqn/NA2000.DAT:AAZ92369 +   75.50  127.31  51.49    1571  !
/SIDS2/gcgdata/geneseq/geneseqn/NA1991.DAT:AAQ13113 +   75.50  122.75  92.34    2541  !
/SIDS2/gcgdata/geneseq/geneseqn/NA1992.DAT:AAQ24235 +   75.50  122.75  92.38    2542  !
/SIDS2/gcgdata/geneseq/geneseqn/NA2000.DAT:AAC75288 +   75.50  122.04  101.20   2740  !
/SIDS2/gcgdata/geneseq/geneseqn/NA2001.DAT:AAS05388 +   75.50  118.58  157.76   3949
/SIDS2/gcgdata/geneseq/geneseqn/NA2001.DAT:AAS23095 +   75.50  117.44  182.49   4452
/SIDS2/gcgdata/geneseq/geneseqn/NA2001.DAT:AAS08694 -   75.00  116.62  202.88   4388
/SIDS2/gcgdata/geneseq/geneseqn/NA1995.DAT:AAQ88228 -   75.00  109.24  522.32   9558
/SIDS2/gcgdata/geneseq/geneseqn/NA2000.DAT:AAF12991 +   74.50  132.52  26.37    739

seq_name: /SIDS2/gcgdata/geneseq/geneseqn/NA1999.DAT:AAX24993

seq_documentation_block:
ID   AAX24993 standard; cDNA; 332 BP.
XX
AC   AAX24993;
XX
DT   05-JUL-1999  (first entry)
XX
DE   Murine Bcl-2 interacting mediator of cell death Bim-S cDNA.
XX
KW   Bim-S; Bcl-2 interacting mediator of cell death; apoptosis;
KW   cell cycle; mouse; cancer; autoimmune disease;
KW   degenerative disease; therapy; contraceptive; splice variant;
KW   isoform; ss.
XX
OS   Mus musculus.
XX
PN   WO9914321-A1.
XX
PD   25-MAR-1999.
XX
PF   17-SEP-1998;   98WO-AU00772.
XX
PR   24-SEP-1997;   97AU-0009373.
PR   17-SEP-1997;   97AU-0009263.
XX
PA   (HALL-) HALL INST MEDICAL RES WALTER & ELIZA.
XX
PI   Adams J,  Cory S,  Huang DCS,  O'Connor L,  O'Reilly L;
PI   Puthalakath H,  Strasser A;
XX
DR   WPI; 1999-244030/20.
DR   P-PSDB; AAW98154.
XX
PT   New isolated member of the Bcl-2 family, Bim used in, e.g. cancer
PT   treatment
XX
PS   Claim 3; Page 92; 145pp; English.
XX
CC   The present sequence encodes the short form (S) of murine Bim, or
CC   Bcl-2 interacting mediator of cell death (see AAW98154), a novel
CC   member of the Bcl-2 family that is capable of inducing cell death
CC   (apoptosis) and which acts as a 'death-ligand' for certain members
CC   of the pro-survival Bcl-2 family.  Bim is a BH3-only protein, as the
CC   only Bcl-2 homology region which it encompasses is BH3.  It is the
CC   only BH3-only protein for which splice variants exist.  These
CC   result in the expression of a variety of isoforms, i.e. Bim-S,
CC   Bim-L and Bim-EL (see AAW98154-56).  cDNAs encoding these murine Bim
CC   isoforms were obtained from a T lymphoma cDNA library using human
CC   recombinant Bcl-2 protein.  The murine Bim gene has been mapped to
CC   chromosome 2 at bands F3-G.  Human Bim-L and Bim-EL isoforms have
CC   also been identified (see AAW98157-58).  Binding the dynein light
CC   chain was shown to regulate the pro-apoptotic activity of Bim.
CC   Bim-S, the splice variant which does not bind to dynein light
CC   chain, is a much more potent killer than either Bim-L or Bim-EL.
CC   The invention provides variants (see AAW98159-68) of murine and human
CC   Bim-L or Bim-EL that cannot bind, couple or otherwise associate
CC   with a dynein light chain.  The identification of Bim permits the
CC   identification and rational design of a range of products for use
CC   in therapy, diagnosis, antibody generation and involving modulation
CC   of physiological cell death.  These therapeutic molecules may act
CC   as either antagonists or agonists of Bim's function and will be
CC   useful in cancer, autoimmune or degenerative disease therapy.
CC   Increased Bim expression or Bim activity is useful, e.g. for
CC   treatment or prophylaxis in conditions such as cancer and deletion
CC   of autoreactive lymphocytes in autoimmune disease.  Decreased Bim
```

```
CC   expression of Bim activity is useful in regulating inhibition or
CC   prevention of cell death or degeneration such as under cytotoxic
CC   conditions during e.g. gamma-irradiation and chemotherapy or during
CC   HIV/AIDS or other viral infections, ischemia, myocardial infarction,
CC   hypoxia, degenerative diseases or for prolonging the survival of
CC   cells being transplanted for treatment of disease.  Since Bim is
CC   expressed in germ cells, modulating Bim expression or Bim activity
CC   is useful, e.g. as a contraceptive or method of sterilization by
CC   preventing generation of fertile sperm.
XX
SQ   Sequence 332 BP; 87 A; 85 C; 91 G; 69 T; 0 other;


alignment_scores:
            Quality:  574.00              Length:    110
              Ratio:  5.218                 Gaps:      0
 Percent Similarity: 100.000   Percent Identity: 100.000

alignment_block:
 US-09-508-832-2 x AAX24993    ..

 Align seg 1/1  to: AAX24993  from: 1  to: 332

       1 MetAlaLysGlnProSerAspValSerSerGluCysAspArgGluGlyGl 17
         ||||||||||||||||||||||||||||||||||||||||||||||||||
       1 ATGGCCAAGCAACCTTCTGATGTAAGTTCTGAGTGTGACAGAGAAGGTGG 50

      17 yGlnLeuGlnProAlaGluArgProProGlnLeuArgProGlyAlaProT 34
         ||||||||||||||||||||||||||||||||||||||||||||||||||
      51 ACAATTGCAGCCTGCTGAGAGGCCTCCCCAGCTCAGGCCTGGGGCCCCTA 100

      34 hrSerLeuGlnThrGluProGlnAlaSerIleArgGlnSerGlnGluGlu 50
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     101 CCTCCCTACAGACAGAACCGCAAGCTTCCATACGACAGTCTCAGGAGGAA 150

      51 ProGluAspLeuArgProGluIleArgIleAlaGlnGluLeuArgArgIl 67
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     151 CCTGAAGATCTGCGCCCGGAGATACGGATTGCACAGGAGCTGCGGCGGAT 200

      67 eGlyAspGluPheAsnGluThrTyrThrArgArgValPheAlaAsnAspT 84
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     201 CGGAGACGAGTTCAACGAAACTTACACAAGGAGGGTGTTTGCAAATGATT 250

      84 yrArgGluAlaGluAspHisProGlnMetValIleLeuGlnLeuLeuArg 100
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     251 ACCGCGAGGCTGAAGACCACCCTCAAATGGTTATCTTACAACTGTTACGC 300

     101 PheIlePheArgLeuValTrpArgArgHis 110
         ||||||||||||||||||||||||||||||
     301 TTTATCTTCCGTCTGGTATGGAGAAGGCAT 330

seq_name: /SIDS2/gcgdata/geneseq/geneseqn/NA1999.DAT:AAX24994

seq_documentation_block:
ID   AAX24994 standard; cDNA; 422 BP.
XX
AC   AAX24994;
XX
DT   05-JUL-1999  (first entry)
XX
DE   Murine Bcl-2 interacting mediator of cell death Bim-L cDNA.
XX
KW   Bim-L; Bcl-2 interacting mediator of cell death; apoptosis;
KW   cell cycle; mouse; cancer; autoimmune disease;
KW   degenerative disease; therapy; contraceptive; splice variant;
KW   isoform; ss.
XX
OS   Mus musculus.
XX
PN   WO9914321-A1.
XX
PD   25-MAR-1999.
XX
PF   17-SEP-1998;  98WO-AU00772.
XX
PR   24-SEP-1997;  97AU-0009373.
PR   17-SEP-1997;  97AU-0009263.
XX
PA   (HALL-) HALL INST MEDICAL RES WALTER & ELIZA.
XX
PI   Adams J,  Cory S,  Huang DCS,  O'Connor L,  O'Reilly L;
PI   Puthalakath H,  Strasser A;
XX
DR   WPI; 1999-244030/20.
DR   P-PSDB; AAW98155.
XX
PT   New isolated member of the Bcl-2 family, Bim used in, e.g. cancer
PT   treatment
XX
PS   Claim 3; Page 94-95; 145pp; English.
XX
CC   The present sequence encodes the long form (L) of murine Bim, or
CC   Bcl-2 interacting mediator of cell death (see AAW98155), a novel
CC   member of the Bcl-2 family that is capable of inducing cell death
CC   (apoptosis) and which acts as a 'death-ligand' for certain members
CC   of the pro-survival Bcl-2 family.  Bim is a BH3-only protein, as the
CC   only Bcl-2 homology region which it encompasses is BH3.  It is the
CC   only BH3-only protein for which splice variants exist.  These
CC   result in the expression of a variety of isoforms, i.e. Bim-S,
CC   Bim-L and Bim-EL (see AAW98154-56).  cDNAs encoding these murine Bim
CC   isoforms were obtained from a T lymphoma cDNA library using human
CC   recombinant Bcl-2 protein.  The murine Bim gene has been mapped to
CC   chromosome 2 at bands F3-G.  Human Bim-L and Bim-EL isoforms have
CC   also been identified (see AAW98157-58).  Binding the dynein light
CC   chain was shown to regulate the pro-apoptotic activity of Bim.
CC   Bim-S, the splice variant which does not bind to dynein light
CC   chain, is a much more potent killer than either Bim-L or Bim-EL.
CC   The invention provides variants (see AAW98159-68) of murine and human
CC   Bim-L or Bim-EL that cannot bind, couple or otherwise associate
CC   with a dynein light chain.  The identification of Bim permits the
CC   identification and rational design of a range of products for use
CC   in therapy, diagnosis, antibody generation and involving modulation
CC   of physiological cell death.  These therapeutic molecules may act
CC   as either antagonists or agonists of Bim's function and will be
CC   useful in cancer, autoimmune or degenerative disease therapy.
CC   Increased Bim expression or Bim activity is useful, e.g. for
CC   treatment or prophylaxis in conditions such as cancer and deletion
CC   of autoreactive lymphocytes in autoimmune disease.  Decreased Bim
CC   expression of Bim activity is useful in regulating inhibition or
CC   prevention of cell death or degeneration such as under cytotoxic
CC   conditions during e.g. gamma-irradiation and chemotherapy or during
CC   HIV/AIDS or other viral infections, ischemia, myocardial infarction,
CC   hypoxia, degenerative diseases or for prolonging the survival of
CC   cells being transplanted for treatment of disease.  Since Bim is
CC   expressed in germ cells, modulating Bim expression or Bim activity
CC   is useful, e.g. as a contraceptive or method of sterilization by
CC   preventing generation of fertile sperm.
XX
SQ   Sequence 422 BP; 112 A; 116 C; 109 G; 85 T; 0 other;


alignment_scores:
            Quality:  549.00              Length:    140
              Ratio:  4.991                 Gaps:      1
 Percent Similarity:  78.571   Percent Identity: 78.571

alignment_block:
 US-09-508-832-2 x AAX24994    ..

 Align seg 1/1  to: AAX24994  from: 1  to: 422

       1 MetAlaLysGlnProSerAspValSerSerGluCysAspArgGluGlyGl 17
         ||||||||||||||||||||||||||||||||||||||||||||||||||
       1 ATGGCCAAGCAACCTTCTGATGTAAGTTCTGAGTGTGACAGAGAAGGTGG 50

      17 yGlnLeuGlnProAlaGluArgProProGlnLeuArgProGlyAlaProT 34
```

```
     ||||||||||||||||||||||||||||||||||||||||||||||||||
  51 ACAATTGCAGCCTGCTGAGAGGCCTCCCCAGCTCAGGCCTGGGGCCCCTA 100

  34 hrSerLeuGlnThrGluProGln..........................  41
     ||||||||||||||||||||||
 101 CCTCCCTACAGACAGAACCGCAAGACAGGAGCCCGGCACCCATGAGTTGT 150

  41 .................................................  41

 151 GACAAGTCAACACAAACCCCAAGTCCTCCTTGCCAGGCCTTCAACCACTA 200

  42 .............AlaSerIleArgGlnSerGlnGluGluProGluAspL  54
                  ||||||||||||||||||||||||||||||||||||||
 201 TCTCAGTGCAATGGCTTCCATACGACAGTCTCAGGAGGAACCTGAAGATC 250

  54 euArgProGluIleArgIleAlaGlnGluLeuArgArgIleGlyAspGlu  70
     ||||||||||||||||||||||||||||||||||||||||||||||||||
 251 TGCGCCCGGAGATACGGATTGCACAGGAGCTGCGGCGGATCGGAGACGAG 300

  71 PheAsnGluThrTyrThrArgArgValPheAlaAsnAspTyrArgGluAl  87
     ||||||||||||||||||||||||||||||||||||||||||||||||||
 301 TTCAACGAAACTTACACAAGGAGGGTGTTTGCAAATGATTACCGCGAGGC 350

  87 aGluAspHisProGlnMetValIleLeuGlnLeuLeuArgPheIlePheA 104
     ||||||||||||||||||||||||||||||||||||||||||||||||||
 351 TGAAGACCACCCTCAAATGGTTATCTTACAACTGTTACGCTTTATCTTCC 400

 104 rgLeuValTrpArgArgHis 110
     ||||||||||||||||||
 401 GTCTGGTATGGAGAAGGCAT 420
```

seq_name: /SIDS2/gcgdata/geneseq/geneseqn/NA1999.DAT:AAX24995

seq_documentation_block:

```
ID   AAX24995 standard; cDNA; 590 BP.
XX
AC   AAX24995;
XX
DT   05-JUL-1999  (first entry)
XX
DE   Murine Bcl-2 interacting mediator of cell death Bim-EL cDNA.
XX
KW   Bim-EL; Bcl-2 interacting mediator of cell death; apoptosis;
KW   cell cycle; mouse; cancer; autoimmune disease;
KW   degenerative disease; therapy; contraceptive; splice variant;
KW   isoform; ss.
XX
OS   Mus musculus.
XX
PN   WO9914321-A1.
XX
PD   25-MAR-1999.
XX
PF   17-SEP-1998;   98WO-AU00772.
XX
PR   24-SEP-1997;   97AU-0009373.
PR   17-SEP-1997;   97AU-0009263.
XX
PA   (HALL-) HALL INST MEDICAL RES WALTER & ELIZA.
XX
PI   Adams J,  Cory S,  Huang DCS,  O'Connor L,  O'Reilly L;
PI   Puthalakath H,  Strasser A;
XX
DR   WPI; 1999-244030/20.
DR   P-PSDB; AAW98156.
XX
PT   New isolated member of the Bcl-2 family, Bim used in, e.g. cancer
PT   treatment
XX
PS   Claim 3; Page 96-97; 145pp; English.
XX
CC   The present sequence encodes the extra long form (EL) of murine Bim,
CC   or Bcl-2 interacting mediator of cell death (see AAW98156), a novel
CC   member of the Bcl-2 family that is capable of inducing cell death
CC   (apoptosis) and which acts as a 'death-ligand' for certain members
CC   of the pro-survival Bcl-2 family.  Bim is a BH3-only protein, as the
CC   only Bcl-2 homology region which it encompasses is BH3.  It is the
CC   only BH3-only protein for which splice variants exist.  These
CC   result in the expression of a variety of isoforms, i.e. Bim-S,
CC   Bim-L and Bim-EL (see AAW98154-56).  cDNAs encoding these murine Bim
CC   isoforms were obtained from a T lymphoma cDNA library using human
CC   recombinant Bcl-2 protein.  The murine Bim gene has been mapped to
CC   chromosome 2 at bands F3-G.  Human Bim-L and Bim-EL isoforms have
CC   also been identified (see AAW98157-58).  Binding the dynein light
CC   chain was shown to regulate the pro-apoptotic activity of Bim.
CC   Bim-S, the splice variant which does not bind to dynein light
CC   chain, is a much more potent killer than either Bim-L or Bim-EL.
CC   The invention provides variants (see AAW98159-68) of murine and human
CC   Bim-L or Bim-EL that cannot bind, couple or otherwise associate
CC   with a dynein light chain.  The identification of Bim permits the
CC   identification and rational design of a range of products for use
CC   in therapy, diagnosis, antibody generation and involving modulation
CC   of physiological cell death.  These therapeutic molecules may act
CC   as either antagonists or agonists of Bim's function and will be
CC   useful in cancer, autoimmune or degenerative disease therapy.
CC   Increased Bim expression or Bim activity is useful, e.g. for
CC   treatment or prophylaxis in conditions such as cancer and deletion
CC   of autoreactive lymphocytes in autoimmune disease.  Decreased Bim
CC   expression of Bim activity is useful in regulating inhibition or
CC   prevention of cell death or degeneration such as under cytotoxic
CC   conditions during e.g. gamma-irradiation and chemotherapy or during
CC   HIV/AIDS or other viral infections, ischemia, myocardial infarction,
CC   hypoxia, degenerative diseases or for prolonging the survival of
CC   cells being transplanted for treatment of disease.  Since Bim is
CC   expressed in germ cells, modulating Bim expression or Bim activity
CC   is useful, e.g. as a contraceptive or method of sterilization by
CC   preventing generation of fertile sperm.
XX
SQ   Sequence 590 BP; 137 A; 178 C; 150 G; 125 T; 0 other;
```

alignment_scores:

```
           Quality:  521.00              Length:    196
             Ratio:  4.736                Gaps:      1
 Percent Similarity:  56.122   Percent Identity: 56.122
```

alignment_block:

US-09-508-832-2 x AAX24995 ..

Align seg 1/1 to: AAX24995 from: 1 to: 590

```
   1 MetAlaLysGlnProSerAspValSerSerGluCysAspArgGluGlyGl  17
     ||||||||||||||||||||||||||||||||||||||||||||||||||
   1 ATGGCCAAGCAACCTTCTGATGTAAGTTCTGAGTGTGACAGAGAAGGTGG  50

  17 yGlnLeuGlnProAlaGluArgProProGlnLeuArgProGlyAlaProT  34
     ||||||||||||||||||||||||||||||||||||||||||||||||||
  51 ACAATTGCAGCCTGCTGAGAGGCCTCCCCAGCTCAGGCCTGGGGCCCCTA 100

  34 hrSerLeuGlnThrGluProGln..........................  41
     ||||||||||||||||||||||
 101 CCTCCCTACAGACAGAACCGCAAGGTAATCCCGACGGCGAAGGGGACCGC 150

  41 .................................................  41

 151 TGCCCCCACGGCAGCCCTCAGGGCCCGCTGGCCCCACCGGCCAGCCCTGG 200

  41 .................................................  41

 201 CCCTTTTGCTACCAGATCCCCACTTTTCATCTTTGTGAGAAGATCTTCTC 250

  41 .................................................  41

 251 TGCTGTCCCGGTCCTCCAGTGGGTATTTCTCTTTTGACACAGACAGGAGC 300

  41 .................................................  41
```

```
     301 CCGGCACCCATGAGTTGTGACAAGTCAACACAAACCCCAAGTCCTCCTTG 350

      42 ..............................AlaSerIleArgGlnSerG 48
                                      ||||||||||||||||||
     351 CCAGGCCTTCAACCACTATCTCAGTGCAATGGCTTCCATACGACAGTCTC 400

      48 lnGluGluProGluAspLeuArgProGluIleArgIleAlaGlnGluLeu 64
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     401 AGGAGGAACCTGAAGATCTGCGCCCGGAGATACGGATTGCACAGGAGCTG 450

      65 ArgArgIleGlyAspGluPheAsnGluThrTyrThrArgArgValPheAl 81
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     451 CGGCGGATCGGAGACGAGTTCAACGAAACTTACACAAGGAGGGTGTTTGC 500

      81 aAsnAspTyrArgGluAlaGluAspHisProGlnMetValIleLeuGlnL 98
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     501 AAATGATTACCGCGAGGCTGAAGACCACCCTCAAATGGTTATCTTACAAC 550

      98 euLeuArgPheIlePheArgLeuValTrpArgArgHis 110
         ||||||||||||||||||||||||||||||||||||
     551 TGTTACGCTTTATCTTCCGTCTGGTATGGAGAAGGCAT 588

seq_name: /SIDS2/gcgdata/geneseq/geneseqn/NA1999.DAT:AAX24996

seq_documentation_block:
ID   AAX24996 standard; cDNA; 416 BP.
XX
AC   AAX24996;
XX
DT   05-JUL-1999  (first entry)
XX
DE   Human Bcl-2 interacting mediator of cell death Bim-L cDNA.
XX
KW   Bim-L; Bcl-2 interacting mediator of cell death; apoptosis;
KW   cell cycle; human; cancer; autoimmune disease;
KW   degenerative disease; therapy; contraceptive; splice variant;
KW   isoform; ss.
XX
OS   Homo sapiens.
XX
PN   WO9914321-A1.
XX
PD   25-MAR-1999.
XX
PF   17-SEP-1998;   98WO-AU00772.
XX
PR   24-SEP-1997;   97AU-0009373.
PR   17-SEP-1997;   97AU-0009263.
XX
PA   (HALL-) HALL INST MEDICAL RES WALTER & ELIZA.
XX
PI   Adams J,  Cory S,  Huang DCS,  O'Connor L,  O'Reilly L;
PI   Puthalakath H,  Strasser A;
XX
DR   WPI; 1999-244030/20.
DR   P-PSDB; AAW98157.
XX
PT   New isolated member of the Bcl-2 family, Bim used in, e.g. cancer
PT   treatment
XX
PS   Claim 7; Page 99-100; 145pp; English.
XX
CC   The present sequence encodes the long form (L) of human Bim, or
CC   Bcl-2 interacting mediator of cell death (see AAW98157), a novel
CC   member of the Bcl-2 family that is capable of inducing cell death
CC   (apoptosis) and which acts as a 'death-ligand' for certain members
CC   of the pro-survival Bcl-2 family.  Bim is a BH3-only protein, as the
CC   only Bcl-2 homology region which it encompasses is BH3.  It is the
CC   only BH3-only protein for which splice variants exist.  These
CC   result in the expression of a variety of isoforms, i.e. Bim-S,
CC   Bim-L and Bim-EL.  cDNAs encoding human Bim-L and Bim-EL (see
CC   AAW98158) were isolated from embryo and liver cDNA libraries using
CC   mouse bim cDNA.  Murine Bim-S, Bim-L and Bim-EL isoforms (see
CC   AAW98154-56) are also provided.  The human Bim gene maps to
CC   chromosome 2 at bands 2q12-2q13.  Binding the dynein light
CC   chain was shown to regulate the pro-apoptotic activity of Bim.
CC   Bim-S, the splice variant which does not bind to dynein light
CC   chain, is a much more potent killer than either Bim-L or Bim-EL.
CC   The invention provides variants (see AAW98159-68) of murine and human
CC   Bim-L or Bim-EL that cannot bind, couple or otherwise associate
CC   with a dynein light chain.  The identification of Bim permits the
CC   identification and rational design of a range of products for use
CC   in therapy, diagnosis, antibody generation and involving modulation
CC   of physiological cell death.  These therapeutic molecules may act
CC   as either antagonists or agonists of Bim's function and will be
CC   useful in cancer, autoimmune or degenerative disease therapy.
CC   Increased Bim expression or Bim activity is useful, e.g. for
CC   treatment or prophylaxis in conditions such as cancer and deletion
CC   of autoreactive lymphocytes in autoimmune disease.  Decreased Bim
CC   expression of Bim activity is useful in regulating inhibition or
CC   prevention of cell death or degeneration such as under cytotoxic
CC   conditions during e.g. gamma-irradiation and chemotherapy or during
CC   HIV/AIDS or other viral infections, ischemia, myocardial infarction,
CC   hypoxia, degenerative diseases or for prolonging the survival of
CC   cells being transplanted for treatment of disease.  Since Bim is
CC   expressed in germ cells, modulating Bim expression or Bim activity
CC   is useful, e.g. as a contraceptive or method of sterilization by
CC   preventing generation of fertile sperm.
XX
SQ   Sequence 416 BP; 113 A; 113 C; 103 G; 87 T; 0 other;


alignment_scores:
            Quality:  436.00              Length:    140
              Ratio:  4.404                 Gaps:      2
 Percent Similarity:  70.714   Percent Identity: 64.286

alignment_block:
 US-09-508-832-2 x AAX24996   ..

 Align seg 1/1  to: AAX24996  from: 1  to: 416

       1 MetAlaLysGlnProSerAspValSerSerGluCysAspArgGluGlyGl 17
         ||||||||||||||||||||||||||||||||||||||||||||||||||
       1 ATGGCAAAGCAACCTTCTGATGTAAGTTCTGAGTGTGACCGAGAAGGTAG 50

      17 yGlnLeuGlnProAlaGluArgProProGlnLeuArgProGlyAlaProT 34
         ||||||||||||||||||||||||||||||||||||||||||||||||||
      51 ACAATTGCAGCCTGCGGAGAGGCCTCCCCAGCTCAGACCTGGGGCCCCTA 100

      34 hrSerLeuGlnThrGluProGln..................................... 41
         |||||||||||||||||||||||
     101 CCTCCCTACAGACAGAGCCACAAGACAGGAGCCCAGCACCCATGAGTTGT 150

      41 ..................................................... 41

     151 GACAAATCAACACAAACCCCAAGTCCTCCTTGCCAGGCCTTCAACCACTA 200

      42 ..............AlaSerIleArgGlnSerGlnGluGluProGluAspL 54
                      ||||||:::||||||:::      ||||||   |||:
     201 TCTCAGTGCAATGGCTTCCATGAGGCAGGCT......GAACCTGCAGATA 244

      54 euArgProGluIleArgIleAlaGlnGluLeuArgArgIleGlyAspGlu 70
         ::|||||||||||||   ||||||||||||||||||||||||||||||||
     245 TGCGCCCAGAGATATGGATCGCCCAAGAGTTGCGGCGTATCGGAGACGAG 294

      71 PheAsnGluThrTyrThrArgArgValPheAlaAsnAspTyrArgGluAl 87
         ||||||     |||:::||||||||||||   |||:::|||:::   ||
     295 TTTAACGCTTACTATGCAAGGAGGGTATTTTTGAATAATTACCAAGCAGC 344

      87 aGluAspHisProGlnMetValIleLeuGlnLeuLeuArgPheIlePheA 104
         ||||||||||||:::||||||||||||:::||||||||:::|||   |
     345 CGAAGACCACCCACGAATGGTTATCTTACGACTGTTACGTTACATTGTCC 394

     104 rgLeuValTrpArgArgHis 110
```

```
          ||||||||||||||   |||
     395 GCCTGGTGTGGAGAATGCAT 414

seq_name: /SIDS2/gcgdata/geneseq/geneseqn/NA1999.DAT:AAX24997

seq_documentation_block:
ID   AAX24997 standard; cDNA; 596 BP.
XX
AC   AAX24997;
XX
DT   05-JUL-1999  (first entry)
XX
DE   Human Bcl-2 interacting mediator of cell death Bim-EL cDNA.
XX
KW   Bim-EL; Bcl-2 interacting mediator of cell death; apoptosis;
KW   cell cycle; human; cancer; autoimmune disease;
KW   degenerative disease; therapy; contraceptive; splice variant;
KW   isoform; ss.
XX
OS   Homo sapiens.
XX
PN   WO9914321-A1.
XX
PD   25-MAR-1999.
XX
PF   17-SEP-1998;   98WO-AU00772.
XX
PR   24-SEP-1997;   97AU-0009373.
PR   17-SEP-1997;   97AU-0009263.
XX
PA   (HALL-) HALL INST MEDICAL RES WALTER & ELIZA.
XX
PI   Adams J,  Cory S,  Huang DCS,  O'Connor L,  O'Reilly L;
PI   Puthalakath H,  Strasser A;
XX
DR   WPI; 1999-244030/20.
DR   P-PSDB; AAW98158.
XX
PT   New isolated member of the Bcl-2 family, Bim used in, e.g. cancer
PT   treatment
XX
PS   Claim 7; Page 101-102; 145pp; English.
XX
CC   The present sequence encodes the extra long form (EL) of human Bim,
CC   or Bcl-2 interacting mediator of cell death (see AAW98158), a novel
CC   member of the Bcl-2 family that is capable of inducing cell death
CC   (apoptosis) and which acts as a 'death-ligand' for certain members
CC   of the pro-survival Bcl-2 family.  Bim is a BH3-only protein, as the
CC   only Bcl-2 homology region which it encompasses is BH3.  It is the
CC   only BH3-only protein for which splice variants exist.  These
CC   result in the expression of a variety of isoforms, i.e. Bim-S,
CC   Bim-L and Bim-EL.  cDNAs encoding human Bim-L and Bim-EL (see
CC   AAW98158) were isolated from embryo and liver cDNA libraries using
CC   mouse bim cDNA.  Murine Bim-S, Bim-L and Bim-EL isoforms (see
CC   AAW98154-56) are also provided.  The human Bim gene maps to
CC   chromosome 2 at bands 2q12-2q13.  Binding the dynein light
CC   chain was shown to regulate the pro-apoptotic activity of Bim.
CC   Bim-S, the splice variant which does not bind to dynein light
CC   chain, is a much more potent killer than either Bim-L or Bim-EL.
CC   The invention provides variants (see AAW98159-68) of murine and human
CC   Bim-L or Bim-EL that cannot bind, couple or otherwise associate
CC   with a dynein light chain.  The identification of Bim permits the
CC   identification and rational design of a range of products for use
CC   in therapy, diagnosis, antibody generation and involving modulation
CC   of physiological cell death.  These therapeutic molecules may act
CC   as either antagonists or agonists of Bim's function and will be
CC   useful in cancer, autoimmune or degenerative disease therapy.
CC   Increased Bim expression or Bim activity is useful, e.g. for
CC   treatment or prophylaxis in conditions such as cancer and deletion
CC   of autoreactive lymphocytes in autoimmune disease.  Decreased Bim
CC   expression of Bim activity is useful in regulating inhibition or
CC   prevention of cell death or degeneration such as under cytotoxic
CC   conditions during e.g. gamma-irradiation and chemotherapy or during
CC   HIV/AIDS or other viral infections, ischemia, myocardial infarction,
CC   hypoxia, degenerative diseases or for prolonging the survival of
CC   cells being transplanted for treatment of disease.  Since Bim is
CC   expressed in germ cells, modulating Bim expression or Bim activity
CC   is useful, e.g. as a contraceptive or method of sterilization by
CC   preventing generation of fertile sperm.
XX
SQ   Sequence 596 BP; 145 A; 175 C; 146 G; 130 T; 0 other;

alignment_scores:
        Quality:  406.00            Length:    200
          Ratio:  4.101               Gaps:      2
 Percent Similarity:  49.500   Percent Identity: 45.000

alignment_block:
 US-09-508-832-2 x AAX24997   ..

 Align seg 1/1  to: AAX24997  from: 1  to: 596

       1 MetAlaLysGlnProSerAspValSerSerGluCysAspArgGluGlyGl 17
         ||||||||||||||||||||||||||||||||||||||||||||||||||
       1 ATGGCAAAGCAACCTTCTGATGTAAGTTCTGAGTGTGACCGAGAAGGTAG 50

      17 yGlnLeuGlnProAlaGluArgProProGlnLeuArgProGlyAlaProT 34
         ||||||||||||||||||||||||||||||||||||||||||||||||||
      51 ACAATTGCAGCCTGCGGAGAGGCCTCCCCAGCTCAGACCTGGGGCCCCTA 100

      34 hrSerLeuGlnThrGluProGln..................................... 41
         |||||||||||||||||||||||
     101 CCTCCCTACAGACAGAGCCACAAGGTAATCCTGAAGGCAATCACGGAGGT 150

      41 ..................................................... 41

     151 GAAGGGGACAGCTGCCCCCACGGCAGCCCTCAGGGCCCGCTGGCCCCACC 200

      41 ..................................................... 41

     201 TGCCAGCCCTGGCCCTTTTGCTACCAGATCCCCGCTTTTCATCTTTATGA 250

      41 ..................................................... 41

     251 GAAGATCCTCCCTGCTGTCTCGATCCTCCAGTGGGTATTTCTCTTTTGAC 300

      41 ..................................................... 41

     301 ACAGACAGGAGCCCAGCACCCATGAGTTGTGACAAATCAACACAAACCCC 350

      42 ..........................................AlaSerI 44
                                                  ||||||:
     351 AAGTCCTCCTTGCCAGGCCTTCAACCACTATCTCAGTGCAATGGCTTCCA 400

      44 leArgGlnSerGlnGluGluProGluAspLeuArgProGluIleArgIle 60
         ::||||||:::      ||||||   |||:::||||||||||||   |||
     401 TGAGGCAGGCT......GAACCTGCAGATATGCGCCCAGAGATATGGATC 444

      61 AlaGlnGluLeuArgArgIleGlyAspGluPheAsnGluThrTyrThrAr 77
         ||||||||||||||||||||||||||||||||||||      |||:::||
     445 GCCCAAGAGTTGCGGCGTATCGGAGACGAGTTTAACGCTTACTATGCAAG 494

      77 gArgValPheAlaAsnAspTyrArgGluAlaGluAspHisProGlnMetV 94
         |||||||||   |||:::|||:::   ||||||||||||||:::||||
     495 GAGGGTATTTTTGAATAATTACCAAGCAGCCGAAGACCACCCACGAATGG 544

      94 alIleLeuGlnLeuLeuArgPheIlePheArgLeuValTrpArgArgHis 110
         ||||||||:::|||||||||:::|||  ||||||||||||||   |||
     545 TTATCTTACGACTGTTACGTTACATTGTCCGCCTGGTGTGGAGAATGCAT 594

seq_name: /SIDS2/gcgdata/geneseq/geneseqn/NA2000.DAT:AAC44501

seq_documentation_block:
ID   AAC44501 standard; DNA; 1747 BP.
XX
```

AC AAC44501;
XX
DT 18-OCT-2000 (first entry)
XX
DE Zea mays DNA fragment SEQ ID NO: 43060.
XX
KW Hybridisation assay; genetic mapping; gene expression control;
KW protein identification; signal transduction pathway; metabolic;
KW pathway; promoter; termination sequence; corn; ss.
XX
OS Zea mays subsp. mays.
XX
PN EP1033405-A2.
XX
PD 06-SEP-2000.
XX
PF 25-FEB-2000; 2000EP-0301439.
XX
PR 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 23-APR-1999; 99US-0130891.
PR 28-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 06-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.

```
PR   26-AUG-1999;   99US-0150884.
PR   27-AUG-1999;   99US-0151065.
PR   27-AUG-1999;   99US-0151066.
PR   27-AUG-1999;   99US-0151080.
PR   30-AUG-1999;   99US-0151303.
PR   31-AUG-1999;   99US-0151438.
PR   01-SEP-1999;   99US-0151930.
PR   07-SEP-1999;   99US-0152363.
PR   10-SEP-1999;   99US-0153070.
PR   13-SEP-1999;   99US-0153758.
PR   15-SEP-1999;   99US-0154018.
PR   16-SEP-1999;   99US-0154039.
PR   20-SEP-1999;   99US-0154779.
PR   22-SEP-1999;   99US-0155139.
PR   23-SEP-1999;   99US-0155486.
PR   24-SEP-1999;   99US-0155659.
PR   28-SEP-1999;   99US-0156458.
PR   29-SEP-1999;   99US-0156596.
PR   04-OCT-1999;   99US-0157117.
PR   05-OCT-1999;   99US-0157753.
PR   06-OCT-1999;   99US-0157865.
PR   07-OCT-1999;   99US-0158029.
PR   08-OCT-1999;   99US-0158232.
PR   12-OCT-1999;   99US-0158369.
PR   13-OCT-1999;   99US-0159293.
PR   13-OCT-1999;   99US-0159294.
PR   13-OCT-1999;   99US-0159295.
PR   14-OCT-1999;   99US-0159329.
PR   14-OCT-1999;   99US-0159330.
PR   14-OCT-1999;   99US-0159331.
PR   14-OCT-1999;   99US-0159637.
PR   14-OCT-1999;   99US-0159638.
PR   18-OCT-1999;   99US-0159584.
PR   21-OCT-1999;   99US-0160741.
PR   21-OCT-1999;   99US-0160767.
PR   21-OCT-1999;   99US-0160768.
PR   21-OCT-1999;   99US-0160770.
PR   21-OCT-1999;   99US-0160814.
PR   21-OCT-1999;   99US-0160815.
PR   22-OCT-1999;   99US-0160980.
PR   22-OCT-1999;   99US-0160981.
PR   22-OCT-1999;   99US-0160989.
PR   25-OCT-1999;   99US-0161404.
PR   25-OCT-1999;   99US-0161405.
PR   25-OCT-1999;   99US-0161406.
PR   26-OCT-1999;   99US-0161359.
PR   26-OCT-1999;   99US-0161360.
PR   26-OCT-1999;   99US-0161361.
PR   28-OCT-1999;   99US-0161920.
PR   28-OCT-1999;   99US-0161992.
PR   28-OCT-1999;   99US-0161993.
PR   29-OCT-1999;   99US-0162142.

alignment_scores:
            Quality:   87.50              Length:    110
              Ratio:  1.458                Gaps:      4
 Percent Similarity:  54.545   Percent Identity: 26.364

alignment_block:
 US-09-508-832-2 x AAC44501   ..

 Align seg 1/1  to: AAC44501  from: 1  to: 1747

       8 ValSerSerGluCysAspArgGluGlyGlyGlnLeuGlnProAlaGluAr 24
         |||      |||   :::::::::||||||           |||::::||
    1094 GTTCCGCGGGAAGATGAGCACCAAGGAGGTGGACGAGCAGATGCTCAACG 1143

      24 gProProGln............LeuArgProGlyAlaProThrSerLeuG 37
         ||||   |||            ||||||         |||   :::   |
    1144 TCCAGAACAAGAACTCGTCCTACTTCGTGGAGTGGATCCCCAACAACGTC 1193

      37 lnThrGluProGlnAlaSerIleArgGlnSerGlnGluGluProGluAsp 53
         ||:::|||                 |||:::   |||:::      |||

    1194 AAGTCGAGCGTGTGCGACATCCCGCCCAGGGGCCTGAAGATGGCGGGGAC 1243

      54 LeuArgProGluIleArgIle.AlaGlnGluLeu...ArgArgIleGlyA 69
         ||||||   |||:::::::::    ||||||:::   ||||||:::::::
    1244 CTTCGTGGGGAACTCCACCTCCATCCAGGAGATGTTCCGCAGGGTGAGCG 1293

      69 spGluPheAsnGluThrTyrThrArgArgValPheAlaAsn..... 82
         :::::|||:::      :::   |||:::::|||   :::
    1294 AGCAGTTCACCGCCATGTTCAGGCGCAAGGCCTTCCTGCACTGGTACACC 1343

      83 .....AspTyrArgGluAlaGluAspHisProGl 92
                :::::   ||||||||||::::::   ::
    1344 GGCGAGGGGATGGATGAGATGGAGTTCACTGAGGCCGAGAGCAACATGAA 1393

      92 nMetValIleLeuGlnLeuLeuArgPhe 101
         :   ::::::   :::      ::::::
    1394 CGATCTGGTCGCCGAGTACCAGCAGTAC 1421

seq_name: /SIDS2/gcgdata/geneseq/geneseqn/NA1994.DAT:AAQ70754

seq_documentation_block:
ID   AAQ70754 standard; DNA; 2206 BP.
XX
AC   AAQ70754;
XX
DT   22-MAR-1995  (first entry)
XX
DE   Beta tubulin gene of Acremonium chrysogenum.
XX
KW   Beta tubulin; Acremonium chrysogenum; transformation;
KW   co-transformation; biosynthesis; cephalosporins; ss.
XX
OS   Acremonium chrysogenum.
XX
FH   Key             Location/Qualifiers
FT   exon            346..357
FT                   /*tag=  a
FT                   /label= Exon 1.
FT   intron          358..519
FT                   /*tag=  b
FT                   /label= Intron 1.
FT   exon            520..543
FT                   /*tag=  c
FT                   /label= Exon 2.
FT   intron          544..610
FT                   /*tag=  d
FT                   /label= Intron 2.
FT   exon            611..733
FT                   /*tag=  e
FT                   /label= Exon 3.
FT   intron          734..807
FT                   /*tag=  f
FT                   /label= Intron 3.
FT   exon            808..1598
FT                   /*tag=  g
FT                   /label= Exon 4.
FT   intron          1599..1661
FT                   /*tag=  h
FT                   /label= Intron 4.
FT   exon            1662..2055
FT                   /*tag=  i
FT                   /label= Exon 5.
XX
PN   EP610842-A.
XX
PD   17-AUG-1994.
XX
PF   07-FEB-1994;   94EP-0101794.
XX
PR   12-FEB-1993;   93DE-4304312.
XX
PA   (FARH ) HOECHST AG.
XX
```

```
PI   Kueck U,  Nowak C,  Walz M;
XX
DR   WPI; 1994-256686/32.
DR   P-PSDB; AAR56512.
XX
PT   New beta-tubulin from Acremonium chrysogenum - and relates DNA,
PT   vectors and transformed cells, for co-transformations of a wide
PT   range of A. chrysogenum strains with foreign genes
XX
PS   Claim 4; Page 9-10; 31pp; German.
XX
CC   Vectors based upon the beta tubulin coding sequence may be used to
CC   transform Acremonium chrysogenum, optionally in combination with
CC   other vectors to introduce a required foreign gene such as the gene
CC   encoding glutaryl acylase, an enzyme involved in cephalosporin
CC   biosynthesis.  The coding sequence may be used to transform a wide
CC   variety of Acremonium chrysogenum strains (wild type and mutants).
CC   Unlike known systems, it is not recipient-strain limited.
XX
SQ   Sequence 2206 BP; 449 A; 740 C; 541 G; 476 T; 0 other;


alignment_scores:
            Quality:   84.50              Length:     95
              Ratio:  1.536                Gaps:      6
 Percent Similarity:  57.895   Percent Identity: 30.526

alignment_block:
 US-09-508-832-2 x AAQ70754    ..

 Align seg 1/1  to: AAQ70754  from: 1  to: 2206

      15 GluGlyGlyGlnLeuGlnProAlaGluArgProProGln..... 27
         ||||||||||:::         |||:::|||||   |||
    1680 GAAGGAGGTCGAGGACCAGATGCGCAACGTCCAGAGCAAGAACTCGTCCT 1729

      28 .LeuArgProGlyAlaProThrSerLeuGlnThrGluProGlnAlaSerI 44
          ||||||         |||         |||   :::|||:::   |||:
    1730 ACTTCGTCGAGTGGATCCC.........CAACAACATCCAGACCGCTCTC 1770

      44 leArgGlnSerGlnGluGluProGluAsp.........LeuArgProGlu 57
         ::|||:::|||:::      |||:::|||         |||:::   :::
    1771 TGCGCCATTCCTCCCCGTGGCCTCAAGATGTCCTCCACCTTCATCGGCAA 1820

      58 IleArgIle.AlaGlnGluLeu...ArgArgIleGlyAspGluPheAsnG 73
         :::::::::    |||||||||   :::|||:::|||:::::::|||:::
    1821 CTCCACCTCCATCCAGGAGCTGTTCAAGCGTGTCGGTGAGCAGTTCACTG 1870

      73 luThrTyrThrArgArgValPheAlaAsn..................... 82
            :::   |||:::::::|||   :::
    1871 CCATGTTCCGTCGCAAGGCTTTCCTGCATTGGTACACTGGTGAGGGCATG 1920

      83 .........AspTyrArgGluAlaGluAspHis 90
                 ::::::   |||||||||:::::::
    1921 GACGAGATGGAGTTTACCGAGGCCGAGTCCAAC 1953

seq_name: /SIDS2/gcgdata/geneseq/geneseqn/NA1993.DAT:AAQ48230

seq_documentation_block:
ID   AAQ48230 standard; DNA; 3445 BP.
XX
AC   AAQ48230;
XX
DT   22-FEB-1994  (first entry)
XX
DE   Acremonium chrysogenum beta-tubulin gene.
XX
KW   Beta tubulin; mutant; chemical resistance; selective marker;
KW   cephalosporin; antibiotic production; ds.
XX
OS   Acremonium chrysogenum.
XX
FH   Key             Location/Qualifiers

FT   exon            1..1298
FT                   /*tag= a
FT                   /number= 1
FT                   /note= "ATG initiation codon is located at
FT                   nucleotides 1287..1289"
FT   intron          1299..1460
FT                   /*tag= b
FT                   /number= 1
FT   exon            1461..1484
FT                   /*tag= c
FT                   /number= 2
FT   intron          1485..1551
FT                   /*tag= d
FT                   /number= 2
FT   exon            1552..1674
FT                   /*tag= e
FT                   /number= 3
FT   intron          1675..1748
FT                   /*tag= f
FT                   /number= 3
FT   exon            1749..2539
FT                   /*tag= g
FT                   /number= 4
FT   intron          2540..2602
FT                   /*tag= h
FT                   /number= 4
FT   exon            2603..3445
FT                   /*tag= i
FT                   /number= 5
FT                   /note= "TAA termination codon is located at
FT                   nucleotides 2994..2996"
XX
PN   JP05192157-A.
XX
PD   03-AUG-1993.
XX
PF   26-MAY-1992;   92JP-0133384.
XX
PR   27-MAY-1991;   91JP-0121276.
XX
PA   (TAKE ) TAKEDA CHEM IND LTD.
XX
DR   WPI; 1993-277472/35.
DR   P-PSDB; AAR40226.
XX
PT   DNA fragment contg. DNA coding mutant beta-tubulin - originates
PT   from Acremonium chrysogenum, used as selective marker for
PT   transformation of A.chrysogenum
XX
PS   Example 5; Fig 4-6; 16pp; Japanese.
XX
CC   The wild-type coding sequence for beta-tubulin was isolated from
CC   Acremonium chrysogenum ATCC 11550 and sequenced (AAQ48230). Primers
CC   CTU-3 and CTU-6 were used to introduce mutations at codon 100 (Asn
CC   to Ile) and 167 (Phe to Tyr), respectively. Expression of the
CC   mutant proteins encoded by these sequences confers chemical
CC   resistance (e.g. to carbendazim and to ansamitocin) on transformed
CC   microorganisms. See AAQ55405 and AAQ55406 for mutated sequences.
XX
SQ   Sequence 3445 BP; 723 A; 1061 C; 892 G; 769 T; 0 other;


alignment_scores:
            Quality:   84.50              Length:     95
              Ratio:  1.536                Gaps:      6
 Percent Similarity:  57.895   Percent Identity: 30.526

alignment_block:
 US-09-508-832-2 x AAQ48230    ..

 Align seg 1/1  to: AAQ48230  from: 1  to: 3445

      15 GluGlyGlyGlnLeuGlnProAlaGluArgProProGln..... 27
```

```
          |||||||||:::          |||:::||||||   |||
     2621 GAAGGAGGTCGAGGACCAGATGCGCAACGTCCAGAGCAAGAACTCGTCCT 2670

       28 .LeuArgProGlyAlaProThrSerLeuGlnThrGluProGlnAlaSerI 44
           ||||||          |||          |||   :::|||:::   |||:
     2671 ACTTCGTCGAGTGGATCCC.........CAACAACATCCAGACCGCTCTC 2711

       44 leArgGlnSerGlnGluGluProGluAsp.........LeuArgProGlu 57
          ::|||:::|||:::       |||:::|||          |||:::   :::
     2712 TGCGCCATTCCTCCCCGTGGCCTCAAGATGTCCTCCACCTTCATCGGCAA 2761

       58 IleArgIle.AlaGlnGluLeu...ArgArgIleGlyAspGluPheAsnG 73
          :::::::::    |||||||||   :::|||:::|||:::::|||:::
     2762 CTCCACCTCCATCCAGGAGCTGTTCAAGCGTGTCGGTGAGCAGTTCACTG 2811

       73 luThrTyrThrArgArgValPheAlaAsn..................... 82
             :::   |||:::::::|||   :::
     2812 CCATGTTCCGTCGCAAGGCTTTCCTGCATTGGTACACTGGTGAGGGCATG 2861

       83 .........AspTyrArgGluAlaGluAspHis 90
                  ::::::   |||||||||:::::::
     2862 GACGAGATGGAGTTTACCGAGGCCGAGTCCAAC 2894

seq_name: /SIDS2/gcgdata/geneseq/geneseqn/NA1993.DAT:AAQ55405

seq_documentation_block:
ID   AAQ55405 standard; DNA; 3445 BP.
XX
AC   AAQ55405;
XX
DT   22-FEB-1994  (first entry)
XX
DE   A.chrysogenum beta-tubulin Ile(100) mutant coding sequence.
XX
KW   Beta tubulin; mutant; chemical resistance; selective marker;
KW   cephalosporin; antibiotic production; ds.
XX
OS   Acremonium chrysogenum.
XX
FH   Key             Location/Qualifiers
FT   exon            1..1298
FT                   /*tag= a
FT                   /number= 1
FT                   /note= "ATG initiation codon is located at
FT                        nucleotides 1287..1289"
FT   intron          1299..1460
FT                   /*tag= b
FT                   /number= 1
FT   exon            1461..1484
FT                   /*tag= c
FT                   /number= 2
FT   intron          1485..1551
FT                   /*tag= d
FT                   /number= 2
FT   exon            1552..1674
FT                   /*tag= e
FT                   /number= 3
FT   intron          1675..1748
FT                   /*tag= f
FT                   /number= 3
FT   exon            1749..2539
FT                   /*tag= g
FT                   /number= 4
FT                   /note= "wild-type AAC (Asn) codon at position
FT                        1887..1889 is mutated to ATC (Ile) codon"
FT   intron          2540..2602
FT                   /*tag= h
FT                   /number= 4
FT   exon            2603..3445
FT                   /*tag= i
FT                   /number= 5
FT                   /note= "TAA termination codon is located at
FT                        nucleotides 2994..2996"
XX
PN   JP05192157-A.
XX
PD   03-AUG-1993.
XX
PF   26-MAY-1992;   92JP-0133384.
XX
PR   27-MAY-1991;   91JP-0121276.
XX
PA   (TAKE ) TAKEDA CHEM IND LTD.
XX
DR   WPI; 1993-277472/35.
DR   P-PSDB; AAR48200.
XX
PT   DNA fragment contg. DNA coding mutant beta-tubulin - originates
PT   from Acremonium chrysogenum, used as selective marker for
PT   transformation of A.chrysogenum
XX
PS   Claim 3 and Example 6; Fig 4-6 and Fig 7; 16pp; Japanese.
XX
CC   The wild-type coding sequence for beta-tubulin was isolated from
CC   Acremonium chrysogenum ATCC 11550 and sequenced (AAQ48230). Primers
CC   CTU-3 and CTU-6 were used to introduce mutations at codon 100 (Asn
CC   to Ile) and 167 (Phe to Tyr), respectively. Expression of the
CC   mutant proteins encoded by these sequences confers chemical
CC   resistance (e.g. to carbendazim and to ansamitocin) on transformed
CC   microorganisms. See also AAQ55406.
XX
SQ   Sequence 3445 BP; 722 A; 1061 C; 892 G; 770 T; 0 other;


alignment_scores:
            Quality:   84.50            Length:     95
              Ratio:  1.536              Gaps:      6
Percent Similarity:  57.895   Percent Identity: 30.526

alignment_block:
 US-09-508-832-2 x AAQ55405   ..

 Align seg 1/1  to: AAQ55405  from: 1  to: 3445

       15 GluGlyGlyGlnLeuGlnProAlaGluArgProProGln..... 27
          |||||||||:::          |||:::||||||   |||
     2621 GAAGGAGGTCGAGGACCAGATGCGCAACGTCCAGAGCAAGAACTCGTCCT 2670

       28 .LeuArgProGlyAlaProThrSerLeuGlnThrGluProGlnAlaSerI 44
           ||||||          |||          |||   :::|||:::   |||:
     2671 ACTTCGTCGAGTGGATCCC.........CAACAACATCCAGACCGCTCTC 2711

       44 leArgGlnSerGlnGluGluProGluAsp.........LeuArgProGlu 57
          ::|||:::|||:::       |||:::|||          |||:::   :::
     2712 TGCGCCATTCCTCCCCGTGGCCTCAAGATGTCCTCCACCTTCATCGGCAA 2761

       58 IleArgIle.AlaGlnGluLeu...ArgArgIleGlyAspGluPheAsnG 73
          :::::::::    |||||||||   :::|||:::|||:::::|||:::
     2762 CTCCACCTCCATCCAGGAGCTGTTCAAGCGTGTCGGTGAGCAGTTCACTG 2811

       73 luThrTyrThrArgArgValPheAlaAsn..................... 82
             :::   |||:::::::|||   :::
     2812 CCATGTTCCGTCGCAAGGCTTTCCTGCATTGGTACACTGGTGAGGGCATG 2861

       83 .........AspTyrArgGluAlaGluAspHis 90
                  ::::::   |||||||||:::::::
     2862 GACGAGATGGAGTTTACCGAGGCCGAGTCCAAC 2894

seq_name: /SIDS2/gcgdata/geneseq/geneseqn/NA1993.DAT:AAQ55406

seq_documentation_block:
ID   AAQ55406 standard; DNA; 3445 BP.
XX
AC   AAQ55406;
XX
DT   22-FEB-1994  (first entry)
```

```
XX
DE   A.chrysogenum beta-tubulin Tyr(167) mutant coding sequence.
XX
KW   Beta tubulin; mutant; chemical resistance; selective marker;
KW   cephalosporin; antibiotic production; ds.
XX
OS   Acremonium chrysogenum.
XX
FH   Key             Location/Qualifiers
FT   exon            1..1298
FT                   /*tag= a
FT                   /number= 1
FT                   /note= "ATG initiation codon is located at
FT                         nucleotides 1287..1289"
FT   intron          1299..1460
FT                   /*tag= b
FT                   /number= 1
FT   exon            1461..1484
FT                   /*tag= c
FT                   /number= 2
FT   intron          1485..1551
FT                   /*tag= d
FT                   /number= 2
FT   exon            1552..1674
FT                   /*tag= e
FT                   /number= 3
FT   intron          1675..1748
FT                   /*tag= f
FT                   /number= 3
FT   exon            1749..2539
FT                   /*tag= g
FT                   /number= 4
FT                   /note= "wild-type TTC (Phe) codon at position
FT                         2088..2090 is mutated to TAC (Tyr) codon"
FT   intron          2540..2602
FT                   /*tag= h
FT                   /number= 4
FT   exon            2603..3445
FT                   /*tag= i
FT                   /number= 5
FT                   /note= "TAA termination codon is located at
FT                         nucleotides 2994..2996"
XX
PN   JP05192157-A.
XX
PD   03-AUG-1993.
XX
PF   26-MAY-1992;   92JP-0133384.
XX
PR   27-MAY-1991;   91JP-0121276.
XX
PA   (TAKE ) TAKEDA CHEM IND LTD.
XX
DR   WPI; 1993-277472/35.
DR   P-PSDB; AAR48201.
XX
PT   DNA fragment contg. DNA coding mutant beta-tubulin - originates
PT   from Acremonium chrysogenum, used as selective marker for
PT   transformation of A.chrysogenum
XX
PS   Claim 3 and Example 6; Fig 4-6 and Fig 7; 16pp; Japanese.
XX
CC   The wild-type coding sequence for beta-tubulin was isolated from
CC   Acremonium chrysogenum ATCC 11550 and sequenced (AAQ48230). Primers
CC   CTU-3 and CTU-6 were used to introduce mutations at codon 100 (Asn
CC   to Ile) and 167 (Phe to Tyr), respectively. Expression of the
CC   mutant proteins encoded by these sequences confers chemical
CC   resistance (e.g. to carbendazim and to ansamitocin) on transformed
CC   microorganisms. See also AAQ55405.
XX
SQ   Sequence 3445 BP; 724 A; 1061 C; 892 G; 768 T; 0 other;
```

```
alignment_scores:
            Quality:   84.50              Length:     95
              Ratio:  1.536                Gaps:      6
 Percent Similarity:  57.895   Percent Identity: 30.526

alignment_block:
 US-09-508-832-2.x AAQ55406   ..

 Align seg 1/1  to: AAQ55406  from: 1  to: 3445

      15 GluGlyGlyGlnLeuGlnProAlaGluArgProProGln..... 27
         ||||||||||:::          |||:::||||||   |||
    2621 GAAGGAGGTCGAGGACCAGATGCGCAACGTCCAGAGCAAGAACTCGTCCT 2670

      28 .LeuArgProGlyAlaProThrSerLeuGlnThrGluProGlnAlaSerI 44
          ||||||        |||         |||   :::|||:::   |||:
    2671 ACTTCGTCGAGTGGATCCC.........CAACAACATCCAGACCGCTCTC 2711

      44 leArgGlnSerGlnGluGluProGluAsp.........LeuArgProGlu 57
         ::|||:::|||:::      |||:::|||         |||:::   :::
    2712 TGCGCCATTCCTCCCCGTGGCCTCAAGATGTCCTCCACCTTCATCGGCAA 2761

      58 IleArgIle.AlaGlnGluLeu...ArgArgIleGlyAspGluPheAsnG 73
         :::::::::    |||||||||   :::|||:::|||:::::::|||:::
    2762 CTCCACCTCCATCCAGGAGCTGTTCAAGCGTGTCGGTGAGCAGTTCACTG 2811

      73 luThrTyrThrArgArgValPheAlaAsn..................... 82
            :::   |||::::::|||   :::
    2812 CCATGTTCCGTCGCAAGGCTTTCCTGCATTGGTACACTGGTGAGGGCATG 2861

      83 .........AspTyrArgGluAlaGluAspHis 90
                  ::::::   |||||||||:::::
    2862 GACGAGATGGAGTTTACCGAGGCCGAGTCCAAC 2894

seq_name: /SIDS2/gcgdata/geneseq/geneseqn/NA2000.DAT:AAA63350

seq_documentation_block:
ID   AAA63350 standard; DNA; 21185 BP.
XX
AC   AAA63350;
XX
DT   06-MAR-2001 (first entry)
XX
DE   Streptomyces globisporus C-1027 gene cluster ORF 25-42.
XX
KW   Enediyne C-1027 biosynthesis gene cluster; apoprotein; chromophore;
KW   cancer; ds.
XX
OS   Streptomyces globisporus.
XX
FH   Key             Location/Qualifiers
FT   CDS             complement (1..632)
FT                   /*tag= a
FT                   /product= "type II NRPS adenylation enzyme"
FT   CDS             1966..4044
FT                   /*tag= b
FT                   /product= "transmembrane transport protein"
FT   CDS             4188..5192
FT                   /*tag= c
FT                   /product= "O-methyl transferase"
FT   CDS             5249..6505
FT                   /*tag= d
FT                   /product= "P450 hydroxylase"
FT   CDS             complement (6628..7734)
FT                   /*tag= e
FT                   /product= "oxidoreductase"
FT   CDS             8370..9410
FT                   /*tag= f
FT                   /product= "ORF 31 protein"
FT   CDS             complement (5539..10361)
FT                   /*tag= g
FT                   /product= "oxidoreductase"
FT   CDS             11261..12094
```

```
FT                   /*tag=  h
FT                   /product= "ORF 33 protein"
FT   CDS             12251..13399
FT                   /*tag=  i
FT                   /product= "ORF 34 protein"
FT   CDS             14046..14900
FT                   /*tag=  j
FT                   /product= "proline oxidase"
FT   CDS             complement (14947..15756)
FT                   /*tag=  k
FT                   /product= "ORF 36 protein"
FT   CDS             complement (15853..16323)
FT                   /*tag=  l
FT                   /product= "ORF 37 protein"
FT   CDS             complement (16460..18115)
FT                   /*tag=  m
FT                   /product= "P450 hydroxylase"
FT   CDS             complement (18112..18642)
FT                   /*tag=  n
FT                   /product= "ORF 39 protein"
FT   CDS             18960..20039
FT                   /*tag=  o
FT                   /product= "ORF 40 protein"
FT   CDS             20065..20919
FT                   /*tag=  p
FT                   /product= "ORF 41 protein"
FT   CDS             complement (20807..21185)
FT                   /*tag=  q
FT                   /product= "ORF 42 protein"
XX
PN   WO200040596-A1.
XX
PD   13-JUL-2000.
XX
PF   06-JAN-2000; 2000WO-US00446.
XX
PR   06-JAN-1999;   99US-0115434.
PR   05-JAN-2000; 2000US-0477962.
XX
PA   (REGC ) UNIV CALIFORNIA.
XX
PI   Shen B,  Liu W,  Christenson SD,  Standage S;
XX
DR   WPI; 2000-465947/40.
DR   P-PSDB; AAB13588, AAB13589, AAB13590, AAB13591, AAB13592, AAB13593,
DR   AAB13594, AAB13595, AAB13596, AAB13597, AAB13598, AAB13600, AAB13601,
DR   AAB13602, AAB13603, AAB13607, AAB13606.
XX
PT   Isolated nucleic acid comprising a nucleic acid encoding any of C-1027
PT   open reading frames (ORFs) -7 to 42, excluding ORF 9 (cagA), useful for
PT   the production of enediyne C-1027 antitumour antibiotics -
XX
PS   Claim 1; Page 130-157; 160pp; English.
XX
CC   The present sequence is the last 21184 bases of the enediyne C-1027 gene
CC   cluster from Streptomyces globisporus. Enediyne C-1027 is an antibiotic,
CC   consisting of an apoprotein and a non-peptidic chromophore, which acts by
CC   damaging DNA. The sequences within the gene cluster, and the proteins
CC   they encode, can be used in the treatment of cancer, along with
CC   antagonists of the protein. Each of the open reading frames is
CC   specifically claimed, excluding ORF 9, which encodes CagA.
XX
SQ   Sequence 21185 BP; 2903 A; 7529 C; 7587 G; 3166 T; 0 other;


alignment_scores:
            Quality:   83.00              Length:     57
              Ratio:  2.243                Gaps:      1
 Percent Similarity:  64.912   Percent Identity: 42.105

alignment_block:
 US-09-508-832-2 x AAA63350/rev   ..
```

```
Align seg 1/1  to reverse of: AAA63350  from: 1  to: 21185

      14 ArgGluGlyGlyGlnLeuGlnProAlaGluArgProProGlnLeuArgPr 30
         |||:::||||||      :::||||||:::|||   |||   :::   ||
   13432 CGGCGCGGTGGACCTCGCGAACCGGCGCGGCGTCAGCCGCCGGTGGCGCC 13383

      30 oGlyAlaProThrSerLeuGlnThrGluProGlnAlaSerIleArgGlnS 47
         |                  :::|||   |||:::|||   ||||||:::|
   13382 G.................CGGACCTTCCCGCGCGCCCGCATCAGGCATT 13351

      47 erGlnGluGluProGluAspLeuArgProGluIleArgIleAlaGlnGlu 63
         ||   :::||||||   |||||||||   |||:::   ::::::
   13350 CCGC.CGCGAGCCGCCAGATCTCCGCGGCGAGGTTGATGTTGGAGTTCTC 13302

      64 LeuArgArgIleGlyAspGlu 70
         ||||||   :::|||:::|||
   13301 CTGCGCCTCCTCGGCCAGGAG 13281

seq_name: /SIDS2/gcgdata/geneseq/geneseqn/NA2000.DAT:AAA63348

seq_documentation_block:
ID   AAA63348 standard; DNA; 63164 BP.
XX
AC   AAA63348;
XX
DT   06-MAR-2001  (first entry)
XX
DE   Streptomyces globisporus C-1027 gene cluster.
XX
KW   Enediyne C-1027 biosynthesis gene cluster; apoprotein; chromophore;
KW   cancer; ds.
XX
OS   Streptomyces globisporus.
XX
FH   Key             Location/Qualifiers
FT   CDS             complement (8..658)
FT                   /*tag=  a
FT                   /product= "ORF -7 protein"
FT   CDS             complement (930..1478)
FT                   /*tag=  b
FT                   /product= "ORF -6 protein"
FT   CDS             complement (1649..2713)
FT                   /*tag=  c
FT                   /product= "ORF -5 protein"
FT   CDS             complement (2850..3237)
FT                   /*tag=  d
FT                   /product= "ORF -4 protein"
FT   CDS             complement (3442..4971)
FT                   /*tag=  e
FT                   /product= "ORF -3 protein"
FT   CDS             5982..7479
FT                   /*tag=  f
FT                   /product= "glycerol phosphate transporter"
FT   CDS             complement (7573..9900)
FT                   /*tag=  g
FT                   /product= "ABC transport/UvrA-like protein"
FT   CDS             complement (9982..11349)
FT                   /*tag=  h
FT                   /product= "Na+/H+ transporter"
FT   CDS             complement (11351..12835)
FT                   /*tag=  i
FT                   /product= "hydroxylase/halogenase"
FT   CDS             13012..14079
FT                   /*tag=  j
FT                   /product= "dNDP-glucose synthase"
FT   CDS             complement (14212..14643)
FT                   /*tag=  k
FT                   /product= "CagA"
FT   CDS             complement (14690..15922)
FT                   /*tag=  l
FT                   /product= "aminotransferase"
FT   CDS             complement (15919..16653)
FT                   /*tag=  m
```

```
FT                   /product= "N-methyl transferase"
FT   CDS             complement (16653..17924)
FT                   /*tag= n
FT                   /product= "C-methyl transferase"
FT   CDS             complement (18031..19191)
FT                   /*tag= o
FT                   /product= "epoxide hydrase"
FT   CDS             complement (19267..19929)
FT                   /*tag= p
FT                   /product= "anthranilate synthase II"
FT   CDS             complement (19926..21407)
FT                   /*tag= q
FT                   /product= "anthranilate synthase I"
FT   CDS             complement (21424..22878)
FT                   /*tag= r
FT                   /product= "coenzyme F390 synthetase"
FT   CDS             complement (22875..23546)
FT                   /*tag= s
FT                   /product= "iron-sulphur flavoprotein"
FT   CDS             complement (23566..24702)
FT                   /*tag= t
FT                   /product= "O-acyl transferase"
FT   CDS             complement (24986..25564)
FT                   /*tag= u
FT                   /product= "epimerase"
FT   CDS             25815..27170
FT                   /*tag= v
FT                   /product= "monooxygenase"
FT   CDS             27214..28593
FT                   /*tag= w
FT                   /product= "glycosyl transferase"
FT   CDS             28590..29588
FT                   /*tag= x
FT                   /product= "dNDP glucose dehydratase"
FT   CDS             29632..31197
FT                   /*tag= y
FT                   /product= "SgcB transmembrane efflux protein"
FT   CDS             31280..32590
FT                   /*tag= z
FT                   /product= "coenzyme F390 synthetase"
FT   CDS             32809..34392
FT                   /*tag= aa
FT                   /product= "hydroxylase"
FT   CDS             complement (34458..35294)
FT                   /*tag= ab
FT                   /product= "ORF 22 protein"
FT                   /transl_except= (pos:35223..35225,aa:Ile)
FT                   /transl_except= (pos:35226..35228,aa:Ala)
FT   CDS             complement (35518..35938)
FT                   /*tag= ac
FT                   /product= "ORF 18 protein"
FT   CDS             36165..37490
FT                   /*tag= ad
FT                   /product= "coenzyme F390 synthetase"
FT   CDS             37559..38938
FT                   /*tag= ae
FT                   /product= "type II NRPS condensation enzyme"
FT   CDS             38983..39264
FT                   /*tag= af
FT                   /product= "type II peptidyl carrier protein"
FT   CDS             complement (39367..40986)
FT                   /*tag= ag
FT                   /product= "aminomutase"
FT   CDS             complement (41052..42611)
FT                   /*tag= ah
FT                   /product= "type II NRPS adenylation enzyme"
FT   CDS             43945..46024
FT                   /*tag= ai
FT                   /product= "transmembrane transport protein"
FT   CDS             46167..47171
FT                   /*tag= aj
FT                   /product= "O-methyl transferase"
FT   CDS             47228..48484
FT                   /*tag= ak
FT                   /product= "P450 hydroxylase"
FT   CDS             complement (48607..49714)
FT                   /*tag= al
FT                   /product= "oxidoreductase"
FT   CDS             50350..51390
FT                   /*tag= am
FT                   /product= "ORF 31 protein"
FT   CDS             complement (51421..52341)
FT                   /*tag= an
FT                   /product= "oxidoreductase"
FT   CDS             53241..54074
FT                   /*tag= ao
FT                   /product= "ORF 33 protein"
FT   CDS             54231..55379
FT                   /*tag= ap
FT                   /product= "ORF 34 protein"
FT   CDS             56026..56880
FT                   /*tag= aq
FT                   /product= "proline oxidase"
FT   CDS             complement (56927..57736)
FT                   /*tag= ar
FT                   /product= "ORF 36 protein"
FT   CDS             complement (57833..58304)
FT                   /*tag= as
FT                   /product= "ORF 37 protein"
FT   CDS             complement (58440..60095)
FT                   /*tag= at
FT                   /product= "P450 hydroxylase"
FT   CDS             complement (60092..60621)
FT                   /*tag= au
FT                   /product= "ORF 39 protein"
FT   CDS             60940..62019
FT                   /*tag= av
FT                   /product= "ORF 40 protein"
FT   CDS             62045..62899
FT                   /*tag= aw
FT                   /product= "ORF 41 protein"
FT   CDS             complement (62787..63164)
FT                   /*tag= ax
FT                   /product= "ORF 42 protein"
XX
PN   WO200040596-A1.
XX
PD   13-JUL-2000.
XX
PF   06-JAN-2000; 2000WO-US00446.
XX
PR   06-JAN-1999;   99US-0115434.
PR   05-JAN-2000; 2000US-0477962.
XX
PA   (REGC ) UNIV CALIFORNIA.
XX
PI   Shen B,  Liu W,  Christenson SD,  Standage S;
XX
DR   WPI; 2000-465947/40.
DR   P-PSDB; AAB13554, AAB13555, AAB13556, AAB13557, AAB13558, AAB13559,
DR   AAB13560, AAB13561, AAB13562, AAB13563, AAB13564, AAB13565, AAB13566,
DR   AAB13567, AAB13568, AAB13569, AAB13570, AAB13571, AAB13572, AAB13573,
DR   AAB13574, AAB13575, AAB13576, AAB13577, AAB13578, AAB13579, AAB13580,
DR   AAB13581, AAB13582, AAB13583, AAB13584, AAB13585, AAB13586, AAB13587,
DR   AAB13588, AAB13589, AAB13590, AAB13591, AAB13592, AAB13593, AAB13594,
DR   AAB13595, AAB13596, AAB13597, AAB13598, AAB13600, AAB13601, AAB13602, AAB13603,
XX
PT   Isolated nucleic acid comprising a nucleic acid encoding any of C-1027
PT   open reading frames (ORFs) -7 to 42, excluding ORF 9 (cagA), useful for
PT   the production of enediyne C-1027 antitumour antibiotics -
XX
PS   Claim 1; Page 78-157; 160pp; English.
XX
CC   The present sequence is the enediyne C-1027 gene cluster from
CC   Streptomyces globisporus. Enediyne C-1027 is an antibiotic, consisting of
CC   an apoprotein and a non-peptidic chromophore, which acts by damaging DNA.
```

```
CC   The sequences within the gene cluster, and the proteins they encode, can
CC   be used in the treatment of cancer, along with antagonists of the

alignment_scores:
            Quality:   83.00             Length:     57
              Ratio:  2.243               Gaps:      1
 Percent Similarity:  64.912   Percent Identity: 42.105

alignment_block:
 US-09-508-832-2 x AAA63348/rev   ..

 Align seg 1/1  to reverse of: AAA63348  from: 1  to: 63164

      14 ArgGluGlyGlyGlnLeuGlnProAlaGluArgProProGlnLeuArgPr 30
         |||:::||||||      :::||||||:::|||   |||   :::    ||
   55411 CGGCGCGGTGGACCTCGCGAACCGGCGCGGCGTCAGCCGCCGGTGGCGCC 55362

      30 oGlyAlaProThrSerLeuGlnThrGluProGlnAlaSerIleArgGlnS 47
         |                 :::|||   |||:::|||   ||||||:::|
   55361 G.................CGGACCTTCCCGCGCGCCCGCATCAGGCATT 55330

      47 erGlnGluGluProGluAspLeuArgProGluIleArgIleAlaGlnGlu 63
         ||   :::||||||   |||||||||   |||:::   ::::::
   55329 CCGC.CGCGAGCCGCCAGATCTCCGCGGCGAGGTTGATGTTGGAGTTCTC 55281

      64 LeuArgArgIleGlyAspGlu 70
         ||||||   :::|||:::|||
   55280 CTGCGCCTCCTCGGCCAGGAG 55260

seq_name: /SIDS2/gcgdata/geneseq/geneseqn/NA2000.DAT:AAZ90014

seq_documentation_block:
ID   AAZ90014 standard; DNA; 1158 BP.
XX
AC   AAZ90014;
XX
DT   05-MAY-2000  (first entry)
XX
DE   Metalloprotease MIFR clone nucleotide sequence.
XX
KW   Metalloprotease in the female reproductive tract; MIFR; human; ds;
KW   matrix metalloprotease; MMP.
XX
OS   Homo sapiens.
XX
PN   JP2000014387-A.
XX
PD   18-JAN-2000.
XX
PF   06-JUL-1998;   98JP-0190869.
XX
PR   06-JUL-1998;   98JP-0190869.
XX
PA   (TAKA/) TAKAHASHI T.
PA   (SDIS-) SDI KK.
XX
DR   WPI; 2000-154341/14.
DR   P-PSDB; AAY78589.
XX
PT   A new metalloprotease and a DNA coding it  -
XX
PS   Example 9; Page 19-20; 21pp; Japanese.
XX
CC   This sequence represents a coding sequence of a human metalloprotease in
CC   the female reproductive tract (MIFR) clone. MIFR is a matrix
CC   metalloprotease (MMP). The invention relates to the metalloprotease
CC   protein which is 390 amino acids in length. A recombinant vector
CC   containing the MIFR gene can be used to create transformants which
CC   produce the metalloprotease in culture.
XX
SQ   Sequence 1158 BP; 163 A; 456 C; 372 G; 167 T; 0 other;

alignment_scores:
            Quality:   79.50             Length:     66
              Ratio:  2.149               Gaps:      3
 Percent Similarity:  56.061   Percent Identity: 39.394

alignment_block:
 US-09-508-832-2 x AAZ90014   ..

 Align seg 1/1  to: AAZ90014  from: 1  to: 1158

       6 SerAspValSerSerGluCysAspArgGluGlyGlyGlnLeuGlnProAl 22
         ||||||||||||           ||||||   :::      ||||||::
     364 AGCGACGTGTCCCCCTTCAGCTTCCGCGAGGTGGCCCCCGAGCAGCCCAG 413

      22 aGlu..............ArgProProGlnLeuArgProGlyAlaProT 34
         ::::              |||||||||      |||||||||   |||:
     414 CGACCTCCGGATAGGTGGGCGCCCGCCCCCGCCCCGGCCCGGC...CCTG 460

      34 hrSerLeuGlnThrGluProGlnAlaSerIleArgGlnSerGlnGluGlu 50
         ::         ::::::|||:::|||      |||         ::::::
     461 CGCGCCCGGCCTCTCAGCCCCGTGCTCCCCCCAGGCTTCTACCCGATCAA 510

      51 ProGluAspLeuArgProGluIleArgIleAlaGlnGluLeuArgArg 66
         |||:::   |||   |||   :::|||   |||      |||||||||
     511 CCACACGGACTG...CCTGGTCTCCGCGCTGCACCACTGCTTCGACGG 555

seq_name: /SIDS2/gcgdata/geneseq/geneseqn/NA2001.DAT:AAH50546

seq_documentation_block:
ID   AAH50546 standard; DNA; 2085 BP.
XX
AC   AAH50546;
XX
DT   21-AUG-2001  (first entry)
XX
DE   Insulin receptor gene exon 1 SEQ ID NO:1.
XX
KW   Insulin receptor; polymorphic site; single nucleotide polymorphism;
KW   SNP; migraine; cephalic pain; insulin receptor signalling pathway;
KW   antimigraine; vulnerary; cluster headache; chronic paroxysmal hemicrania;
KW   vascular disorder associated headache; withdrawal; tension headache; ds.
XX
OS   Homo sapiens.
XX
PN   WO200128539-A2.
XX
PD   26-APR-2001.
XX
PF   19-OCT-2000; 2000WO-GB04031.
XX
PR   19-OCT-1999;   99GB-0024713.
PR   19-OCT-1999;   99US-0160423.
XX
PA   (GLAX ) GLAXO GROUP LTD.
XX
PI   Purvis IJ,  McCarthy LC;
XX
DR   WPI; 2001-300274/31.
XX
PT   Use of agent that modulates directly or indirectly insulin receptor or
PT   insulin receptor signaling pathway in the manufacture of medicament for
PT   preventing or treating cephalic pain  -
XX
PS   Disclosure; Page 39-40; 58pp; English.
XX
CC   The present invention describes the use of an agent (I) that modulates
CC   directly or indirectly the insulin receptor or insulin receptor
CC   signalling pathway in the manufacture of a medicament for preventing or
CC   treating cephalic pain. Also described is an isolated polynucleotide (II)
CC   or protein (III) comprising a polymorphism that causes susceptibility to
CC   cephalic pain, or a naturally occurring polymorphism that is in linkage
CC   disequilibrium with the first polymorphism. (I) has antimigraine and
CC   vulnerary activities. (I) is useful for treating cephalic pain which may
```

```
CC   be a cluster headache, chronic paroxysmal hemicrania, headache
CC   associated with vascular disorders, headache associated with substances
CC   or their withdrawal (for example drug withdrawal), tension headache and
CC   in particular migraine with or without aura. The present sequence
CC   represents an exon from the insulin receptor gene, which is used in the
CC   exemplification of the present invention.
XX
SQ   Sequence 2085 BP; 417 A; 631 C; 702 G; 335 T; 0 other;


alignment_scores:
           Quality:   79.00            Length:     63
             Ratio:  2.135              Gaps:      3
 Percent Similarity:  58.730   Percent Identity: 34.921

alignment_block:
 US-09-508-832-2 x AAH50546   ..

 Align seg 1/1  to: AAH50546  from: 1  to: 2085

       2 AlaLysGlnProSerAspValSerSerGluCysAspArgGluGlyGlyGl 18
         |||:::   |||:::|||   :::   :::|||   |||   |||:::
    1329 GCACGGGGCCCAGCTGAC...GGGCCGCGTTGTTTACGGGCCGGAGCAGC 1375

      18 nLeuGlnProAlaGluArgProProGlnLeuArgProGlyAla......P 33
         |||:::||||||   |||||||||         |||||||||      |
    1376 CCTCTCTCCCGCCGCCCGCCCGCCACCCGCCAGCCCAGGTGCCCGCCCGC 1425

      33 roThrSerLeuGlnThrGluProGlnAlaSerIleArgGlnSerGlnGlu 49
         ||:::|||   ::::::   |||::::::   :::   :::|||   :::
    1426 CAGTCAGCTAGTCCGTCGGTCCGCGCGTCCCTCTGTCCCGGAGCCCGCAG 1475

      50 ........................GluProGluAsp 53
                                 ||||||:::|||
    1476 ATCGCGACCCAGAGCGCGCGGGGCCGAGAGCCGAGAGAC 1514

seq_name: /SIDS2/gcgdata/geneseq/geneseqn/NA2001.DAT:AAH31146

seq_documentation_block:
ID   AAH31146 standard; DNA; 2085 BP.
XX
AC   AAH31146;
XX
DT   25-JUL-2001  (first entry)
XX
DE   Human insulin receptor gene exon 1.
XX
KW   Human; insulin receptor; diagnosis; cephalic pain; susceptibility;
KW   single nucleotide polymorphism; SNP; migraine; SNP detection; ds.
XX
OS   Homo sapiens.
XX
PN   WO200129255-A2.
XX
PD   26-APR-2001.
XX
PF   19-OCT-2000; 2000WO-GB04024.
XX
PR   19-OCT-1999;   99GB-0024717.
XX
PA   (GLAX ) GLAXO GROUP LTD.
XX
PI   Purvis IJ,  McCarthy LC;
XX
DR   WPI; 2001-328499/34.
XX
PT   Diagnosis of susceptibility to migraine in individuals comprises the
PT   identification of polymorphisms in the insulin receptor gene region
XX
PS   Disclosure; Page 24-25; 41pp; English.
XX
CC   The present sequence is provided in a specification relating to a
CC   method of diagnosing susceptibility to cephalic pain. Polymorphisms
CC   in the insulin receptor gene have been found to cause susceptibility to
CC   cephalic pain, particularly to migraine. The method comprises typing
CC   a region of the insulin receptor gene or insulin receptor protein of
CC   an individual. A susceptibility to cephalic pain may be diagnosed
CC   using a probe, primer or antibody which is capable of detecting a
CC   polymorphism in an insulin receptor gene region or the insulin
CC   receptor protein.
XX
SQ   Sequence 2085 BP; 417 A; 631 C; 702 G; 335 T; 0 other;


alignment_scores:
           Quality:   79.00            Length:     63
             Ratio:  2.135              Gaps:      3
 Percent Similarity:  58.730   Percent Identity: 34.921

alignment_block:
 US-09-508-832-2 x AAH31146   ..

 Align seg 1/1  to: AAH31146  from: 1  to: 2085

       2 AlaLysGlnProSerAspValSerSerGluCysAspArgGluGlyGlyGl 18
         |||:::   |||:::|||   :::   :::|||   |||   |||:::
    1329 GCACGGGGCCCAGCTGAC...GGGCCGCGTTGTTTACGGGCCGGAGCAGC 1375

      18 nLeuGlnProAlaGluArgProProGlnLeuArgProGlyAla......P 33
         |||:::||||||   |||||||||         |||||||||      |
    1376 CCTCTCTCCCGCCGCCCGCCCGCCACCCGCCAGCCCAGGTGCCCGCCCGC 1425

      33 roThrSerLeuGlnThrGluProGlnAlaSerIleArgGlnSerGlnGlu 49
         ||:::|||   ::::::   |||::::::   :::   :::|||   :::
    1426 CAGTCAGCTAGTCCGTCGGTCCGCGCGTCCCTCTGTCCCGGAGCCCGCAG 1475

      50 ........................GluProGluAsp 53
                                 ||||||:::|||
    1476 ATCGCGACCCAGAGCGCGCGGGGCCGAGAGCCGAGAGAC 1514
```

```
OM of: US-09-508-832-2 to: EST:*   out_format : pfs

Date: Dec 11, 2001  1:03 AM

About: Results were produced by the GenCore software, version 4.5,
       Copyright (c) 1993-2000 Compugen Ltd.

Command line parameters:
 -MODEL=frame+_p2n.model -DEV=xlp
 -Q=/cgn2_1/USPTO_spool/US09508832/runat_10122001_110349_29536/app_query.fasta_1.620
 -DB=EST -QFMT=fastap -SUFFIX=rst -GAPOP=12.000 -GAPEXT=4.000
 -MINMATCH=0.100 -LOOPCL=0.000 -LOOPEXT=0.000 -QGAPOP=4.500
 -QGAPEXT=0.050 -XGAPOP=10.000 -XGAPEXT=0.500 -FGAPOP=6.000
 -FGAPEXT=7.000 -YGAPOP=10.000 -YGAPEXT=0.500 -DELOP=6.000
 -DELEXT=7.000 -START=1 -MATRIX=blosum62 -TRANS=human40.cdi
 -LIST=45 -DOCALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0
 -ALIGN=15 -MODE=LOCAL -OUTFMT=pfs -NORM=ext -MINLEN=0
 -MAXLEN=2000000000 -USER=US09508832_@CGN1_1_0 -NCPU=6 -ICPU=3
 -LONGLOG -NO_XLPXY -WAIT -THREADS=1

Search information block:
Query: US-09-508-832-2
Query length: 110
Database: EST:*
Database sequences: 11351937
Database length: 1077921985
Search time (sec): 2629.110000


score_list:
Sequence              Strd Orig    ZScore    EScore Len   ! Documentation  ..
gb_htc:AK011490         +  493.50  833.29  3.0e-37   1206  ! AK011490 Mus musculus 10 days
gb_est2:BG921698        +  435.00  735.82  8.0e-32    935  ! BG921698 602825518F1 NCI_CGAP_M
gb_est2:BF021882        +  357.00  609.30  9.0e-25    452  ! BF021882 uy59b09.y1 McCarrey Ed
gb_est2:BF319454        -  249.00  426.51  1.4e-14    389  ! BF319454 uy59b09.x1 McCarrey Ed
gb_gss:AZ706148         +  218.00  370.12  1.9e-11    580  ! AZ706148 RPCI-23-227P3.TV RPCI-
gb_est2:BG173095        +  218.00  368.86  2.2e-11    668  ! BG173095 602336666F1 NCI_CGAP_M
gb_est1:AI971169        -  209.00  356.24  1.1e-10    492  ! AI971169 wr24h12.x1 NCI_CGAP_Pr
gb_est1:AA629308        -  127.00  216.28  0.0070     501  ! AA629308 zu84g06.s1 Soares_test
gb_est2:BF172831        +  126.00  222.30  0.0032     210  ! BF172831 PCL5805 Myeloma (PCL)
gb_est1:AA629050        -  123.00  209.46  0.0167     501  ! AA629050 zu84a06.s1 Soares_test
gb_est1:AW629314        -  123.00  206.96  0.0231     664  ! AW629314 hi56e02.x1 Soares_NFL_
gb_est1:AF209718        +   98.00  166.22  4.28       537  ! AF209718 AF209718 Xenopus laevi
gb_est2:BG990772        -   92.00  161.50  7.85       289  ! BG990772 MR1-HT1136-260101-015-
gb_gss:CNS0085G         +   91.00  150.09  33.92      862  ! AL051751 Drosophila melanogaste
gb_est2:BG420514        +   89.50  146.94  50.78      921  ! BG420514 602452566F1 NIH_MGC_14
gb_gss:CNS01YJV         -   89.00  147.57  46.88      780  ! AL173092 Tetraodon nigroviridis
gb_est1:AW432805        +   88.00  149.67  35.79      508  ! AW432805 sh81c06.y1 Gm-c1016 Gl
gb_est2:BG356986        +   86.50  148.93  39.35      414  ! BG356986 OV2_12_D12.g1_A002 Ova
gb_est1:BE415908        +   86.00  146.73  52.20      482  ! BE415908 MUG002.C09R990520 ITEC
gb_est1:BE416016        +   86.00  144.78  66.99      600  ! BE416016 MUG003.D09R990623 ITEC
gb_est1:BE415913        +   86.00  141.32  104.45     886  ! BE415913 MUG002.D02R990520 ITEC
gb_est1:AL558285        +   85.50  140.19  120.72     914  ! AL558285 AL558285 LTI_NFL008_TC
gb_est1:BE416351        +   85.00  144.70  67.73      500  ! BE416351 MUG007.G07R990628 ITEC
gb_gss:CNS00320         -   84.50  137.70  166.23     999  ! AL063754 Drosophila melanogaste
gb_est2:BF859520        +   84.50  131.30  377.46     2052  ! BF859520 963002G02.y4 C. reinh
gb_est2:BG335529        +   84.00  138.73  145.62     808  ! BG335529 602403821F1 NIH_MGC_21
gb_est1:AI970428        -   83.00  155.85  16.19      97  ! AI970428 wr10d03.x1 NCI_CGAP_Lul
gb_est2:BF292486        +   83.00  141.61  100.59     482  ! BF292486 WHE2214_D03_G06ZS Aegi
gb_est1:AI166767        +   83.00  141.58  101.07     484  ! AI166767 xylem.est.567 Poplar x
gb_est2:BF584857        -   83.00  138.97  141.17     649  ! BF584857 602098724F1 NCI_CGAP_C
gb_gss:AZ347259         -   83.00  138.48  150.38     686  ! AZ347259 1M0083H18F Mouse 10kb
gb_est1:AA132780        +   82.00  141.59  100.93     399  ! AA132780 zo18a01.r1 Stratagene
gb_est1:AA099932        +   82.00  141.41  103.24     407  ! AA099932 zl79c04.r1 Stratagene
gb_est1:AA130795        +   82.00  140.00  123.70     477  ! AA130795 zo13g06.r1 Stratagene
gb_est1:AA131161        +   82.00  138.57  148.50     560  ! AA131161 zo16h07.r1 Stratagene
gb_gss:CNS04TRG         -   82.00  135.07  232.82     831  ! AL306853 Tetraodon nigroviridis
gb_gss:CNS048M2         +   82.00  134.54  249.16     882  ! AL279443 Tetraodon nigroviridis
gb_est2:BE915234        +   82.00  134.29  257.23     907  ! BE915234 601667401F1 NCI_CGAP_M
gb_est2:W91244          +   81.50  140.51  115.81     409  ! W91244 mf72g11.r1 Soares mouse
gb_est1:AA575286        +   81.50  139.86  125.86     440  ! AA575286 vi11b09.r1 Barstead mo
gb_est1:AA028436        +   81.50  139.07  139.31     481  ! AA028436 mi19h06.r1 Soares mous
gb_est1:AA008848        +   81.50  138.80  144.27     496  ! AA008848 mg98f07.r1 Soares m
gb_est1:AA015247        +   81.50  138.71  145.92     501  ! AA015247 mh21c06.r1 Soares m
gb_est2:BI182623        -   81.50  135.26  227.22     739  ! BI182623 UNL-P-FN-bm-c-09-0-
gb_est2:BF143712        +   81.50  135.13  231.07     750  ! BF143712 601789785F1 NCI_CGA

seq_name: gb_htc:AK011490

seq_documentation_block:
LOCUS       AK011490     1206 bp    mRNA            HTC       05-JUL-2001
DEFINITION  Mus musculus 10 days embryo cDNA, RIKEN full-length enriched
            library, clone:2610020M23, full insert sequence.
ACCESSION   AK011490
VERSION     AK011490.1  GI:12847647
KEYWORDS    CAP trapper.
SOURCE      Mus musculus (strain:C57BL/6J) 10 days embryo cDNA to mRNA,
            clone_lib:RIKEN full-length enriched mouse cDNA library
            clone:2610020M23.
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 1206)
  AUTHORS   Carninci,P. and Hayashizaki,Y.
  TITLE     High-efficiency full-length cDNA cloning
  JOURNAL   Methods in enzymology. 303, 19-44 (1999)
  MEDLINE   99279253
   PUBMED   10349636
REFERENCE   2  (bases 1 to 1206)
  AUTHORS   Carninci,P., Shibata,Y., Hayatsu,N., Sugahara,Y., Shibata,K.,
            Itoh,M., Konno,H., Okazaki,Y., Muramatsu,M. and Hayashizaki,Y.
  TITLE     Normalization and subtraction of cap-trapper-selected cDNAs to
            prepare full-length cDNA libraries for rapid discovery of new genes
  JOURNAL   Genome research. 10 (10), 1617-1630 (2000)
  MEDLINE   20499374
   PUBMED   11042159
REFERENCE   3  (bases 1 to 1206)
  AUTHORS   Shibata,K., Itoh,M., Aizawa,K., Nagaoka,S., Sasaki,N., Carninci,P.,
            Konno,H., Akiyama,J., Nishi,K., Kitsunai,T., Tashiro,H., Itoh,M.,
            Sumi,N., Ishii,Y., Nakamura,S., Hazama,M., Nishine,T., Harada,A.,
            Yamamoto,R., Matsumoto,H., Sakaguchi,S., Ikegami,T., Kashiwagi,K.,
            Fujiwake,S., Inoue,K., Togawa,Y., Izawa,M., Ohara,E., Watahiki,M.,
            Yoneda,Y., Ishikawa,T., Ozawa,K., Tanaka,T., Matsuura,S., Kawai,J.,
            Okazaki,Y., Muramatsu,M., Inoue,Y., Kira,A. and Hayashizaki,Y.
  TITLE     RIKEN integrated sequence analysis (RISA) system--384-format
            sequencing pipeline with 384 multicapillary sequencer
  JOURNAL   Genome research. 10 (11), 1757-1771 (2000)
  MEDLINE   20530913
   PUBMED   11076861
REFERENCE   4  (bases 1 to 1206)
  AUTHORS   The RIKEN Genome Exploration Research Group Phase II Team and the
            FANTOM Consortium.
  TITLE     Functional annotation of a full-length mouse cDNA collection
  JOURNAL   Nature 409, 685-690 (2001)
REFERENCE   5  (bases 1 to 1206)
  AUTHORS   Adachi,J., Aizawa,K., Akahira,S., Akimura,T., Aono,H., Arai,A.,
            Arakawa,T., Carninci,P., Fukuda,S., Fukunishi,Y., Furuno,M.,
            Hanagaki,T., Hara,A., Hayatsu,N., Hiramoto,K., Hiraoka,T., Hori,F.,
            Imotani,K., Ishii,Y., Itoh,M., Izawa,M., Kato,H., Kawai,J.,
            Kojima,Y., Konno,H., Kouda,M., Koya,S., Kurihara,C., Matsuyama,T.,
            Miyazaki,A., Nishi,K., Nomura,K., Numazaki,R., Ohno,M., Okazaki,Y.,
            Okido,T., Owa,C., Saito,H., Saito,R., Sakai,C., Sakai,K., Sano,H.,
            Sasaki,D., Shibata,K., Shibata,Y., Shinagawa,A., Shiraki,T.,
            Sogabe,Y., Suzuki,H., Tagami,M., Tagawa,A., Takahashi,F.,
            Tanaka,T., Tejima,Y., Toya,T., Yamamura,T., Yasunishi,A.,
            Yoshida,K., Yoshino,M., Muramatsu,M. and Hayashizaki,Y.
  TITLE     Direct Submission
  JOURNAL   Submitted (10-JUL-2000) Yoshihide Hayashizaki, The Institute of
            Physical and Chemical Research (RIKEN), Laboratory for Genome
            Exploration Research Group, RIKEN Gemomic Sciences Center (GSC),
            RIKEN Yokohama Institute; 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama,
            Kanagawa 230-0045, Japan (E-mail:genome-res@gsc.riken.go.jp,
            URL:http://genome.gsc.riken.go.jp/, Tel:81-45-503-9222,
            Fax:81-45-503-9216)
COMMENT     Please visit our web site (http://genome.gsc.riken.go.jp/) for
```

further details.
cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues. First strand cDNA was primed with a primer [5' GAGAGAGAGAAGGATCCAAGAGCTCTTTTTTTTTTTTTTVN 3'], cDNA was prepared by using trehalose thermo-activated reverse transcriptase and subsequently enriched for full-length by cap-trapper. cDNA went through one round of normalization to Rot = 7.5 and subtraction to Rot = 37.5. Second strand cDNA was prepared with the primer adapter of sequence [5' GAGAGAGAGATTCTCGAGTTAATTAAATTAATCCCCCCCCCCCC 3']. cDNA was cleaved with XhoI and SstI. Cloning sites, 5' end: XhoI; 3' end: SstI. Host: SOLR.

```
FEATURES             Location/Qualifiers
     source          1. .1206
                     /organism="Mus musculus"
                     /strain="C57BL/6J"
                     /db_xref="taxon:10090"
                     /db_xref="MGD:MGI:1197519"
                     /db_xref="MGD:MGI:1902115"
                     /clone="2610020M23"
                     /clone_lib="RIKEN full-length enriched mouse cDNA library"
                     /dev_stage="10 days embryo"
BASE COUNT      265 a    339 c    298 g    301 t      3 others
ORIGIN


alignment_scores:
           Quality:  493.50            Length:    196
             Ratio:  4.569               Gaps:      2
 Percent Similarity:  55.102   Percent Identity: 54.592

alignment_block:
 US-09-508-832-2 x AK011490   ..

 Align_seg 1/1  to: AK011490  from: 1  to: 1206

       1 MetAlaLysGlnProSerAspValSerSerGluCysAspArgGluGlyGl 17
         |||||||||||||||||||||||   ||||||||||||   ||||||||
     220 ATGGCCAAGCAACCTTCTGATGTACGTTCTGAGTGTGACAGCGAAGGTGG 269

      17 yGlnLeuGlnProAlaGluArgProProGlnLeuArgProGlyAlaProT 34
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     270 ACAATTGCAGCCTGCTGAGAGGCCTCCCCAGCTCAGGCCTGGGGCCCCTA 319

      34 hrSerLeuGlnThrGluProGln..................................... 41
         ||||||||||||||||||||||
     320 CCTCCCTACAGACAGAACCGCAACGTAATCCGACGGGGAAGGGGACCGTG 369

      41 ..................................................... 41

     370 CCCCCACGGCAGCCCTCAGGGCCCGCTGGCCCCACCGGCCAGCCCTGGCC 419

      41 ..................................................... 41

     420 CTTTTGCTACCAGATCCCCACTTTTCATCTTTGTGAGAAGATCTTCTCTG 469

      41 ..................................................... 41

     470 CTGTCCCGGTCCTCCAGTGGGTATTTCTCTTTTGACACAGACAGGAGCCC 519

      41 ..................................................... 41

     520 GGCACCCATGAGTTGTGACAAGTCAACACAAACCCCAAGTCCTCCTTGCC 569

      42 ............................Ala.SerIleArgGlnSerGln 48
                                     ::: ||||||||||||||||||
     570 AGGCCTTCAACCACTATCTCAGTGCAATGGCTTCCATACGACAGTCTCAG 619

      49 GluGluProGluAspLeuArgProGluIleArgIleAlaGlnGluLeuAr 65
         ||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
     620 GAGGAACCTGAAGATCTGCGCCCGGAGATACGGATTGCACAGGAGCTGCG 669

      65 gArgIleGlyAspGluPheAsnGluThrTyrThrArgArgValPheAlaA 82
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     670 GCGGATCGGAGACGAGTTCAACGAAACTTACACAAGGAGGGTGTTTGCAA 719

      82 snAspTyrArgGluAlaGluAspHisProGlnMetValIleLeuGlnLeu 98
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     720 ATGATTACCGCGAGGCTGAAGACCACCCTCAAATGGTTATCTTACAACTG 769

      99 LeuArgPheIlePheArgLeuValTrpArgArgHis 110
         |||||||||||||||||||||||||||||||||||
     770 TTACGCTTTATCTTCCGTCTGGTATGGAGAAGGCAT 805

seq_name: gb_est2:BG921698

seq_documentation_block:
LOCUS       BG921698      935 bp    mRNA            EST       05-JUN-2001
DEFINITION  602825518F1 NCI_CGAP_Mam6 Mus musculus cDNA clone IMAGE:4954300 5',
            mRNA sequence.
ACCESSION   BG921698
VERSION     BG921698.1  GI:14302174
KEYWORDS    EST.
SOURCE      house mouse.
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 935)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Jeffrey Green M.D.
            cDNA Library Preparation: Life Technologies, Inc.
            cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
            DNA Sequencing by: Incyte Genomics, Inc.
            Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLAM10915  row: c  column: 05
            High quality sequence start: 3
            High quality sequence stop: 786.
FEATURES             Location/Qualifiers
     source          1. .935
                     /organism="Mus musculus"
                     /strain="FVB/N"
                     /db_xref="taxon:10090"
                     /clone="IMAGE:4954300"
                     /clone_lib="NCI_CGAP_Mam6"
                     /sex="female, virgin"
                     /tissue_type="infiltrating ductal carcinoma"
                     /dev_stage="5 months"
                     /lab_host="DH10B"
                     /note="Organ: mammary; Vector: pCMV-SPORT6; Site_1: SalI;
                     Site_2: NotI; Cloned unidirectionally.  Primer: Oligo dT.
                     Library constructed by Life Technologies. Investigator
                     providing samples: Jeffrey Green, M.D., NIH"
BASE COUNT      203 a    283 c    276 g    173 t
ORIGIN


alignment_scores:
           Quality:  435.00            Length:    197
             Ratio:  4.028               Gaps:      3
 Percent Similarity:  54.822   Percent Identity: 53.299

alignment_block:
 US-09-508-832-2 x BG921698   ..

 Align seg 1/1  to: BG921698  from: 1  to: 935

       1 MetAlaLysGlnProSerAspValSerSerGluCysAspArgGluGlyGl 17
```

```
          ||||||||||||||||||||||||||||||||||||||||||||||||||
      208 ATGGCCAAGCAACCTTCTGATGTAAGTTCTGAGTGTGACAGAGAAGGTGG 257

       17 yGlnLeuGlnProAlaGluArgProProGlnLeuArgProGlyAlaProT 34
          ||||||||||||||||||||||||||||||||||||||||||||||||||
      258 ACAATTGCAGCCTGCTGAGAGGCCTCCCCAGCTCAGGCCTGGGGCCCCTA 307

       34 hrSerLeuGlnThrGluPro.Gln..................... 41
          |||||||||||||||||||  :::
      308 CCTCCCTACAGACAGAACCGCAAGGTAATCCCGACGGCGAAGGGGACCGC 357

       41 ............................................. 41

      358 TGCCCCCACGGCATGCCCTCAGGGCCCGCTGGCCCCACCGGCCAGCCCTG 407

       41 ............................................. 41

      408 GCCCTTTTTGCTACCAGATCCCCACTTTTCATCTTTGTGAGAAGATCTTC 457

       41 ............................................. 41

      458 TCTGCTGTCCCGGTCCTCCAGTGGGTATTTCTCTTTTGACACAGACAGGA 507

       41 ............................................. 41

      508 GCCCGGCACCCATGAGTTGTGACAAGTCAACACAAACCCAAGTCCTCCTT 557

       42 ............................AlaSerIleArgGlnSer 47
                                      ||||||||||||||||||
      558 GCCAGGCCTTCAACCACTATCTCAGTGCAATGGCTTCCATACGACAGTCT 607

       48 GlnGluGluProGluAspLeuArgProGluIleArgIleAlaGlnGluLe 64
          ||||||||||||||||||||||||||||||||||||||||||||||||||
      608 CAGGAGGAACCTGAAGATCTGCGCCCGGAGATACGGATTGCACAGGAGCT 657

       64 uArgArgIleGlyAspGluPheAsnGluThrTyrThrArgArgValPheA 81
          ||||||||||||||||||||||||||||||||||||||||||||||||||
      658 GCGGCGGATCGGAGACGAGTTCAACGAAACTTACACAAGGAGGGTGTTTG 707

       81 laAsnAspTyrArgGluAlaGluAspHisProGlnMetValIleLeuGln 97
          ||||||||||||||||||||||   ||||||||||||||||||||||||
      708 CAAATGATTACCGCGAGGCTGAAGA.CACCCTCAAATGGTTATCTTACAA 756

       98 ..LeuLeuArgPheIlePheArgLeuValTrpArgArg 109
            :::|||||||||||||||||||||||||||   :::
      757 ACGTGTTACGCTTTATCTTCCGTCTGGTATGGGCGAAA 794
```

seq_name: gb_est2:BF021882

seq_documentation_block:

```
LOCUS       BF021882      452 bp    mRNA            EST       29-DEC-2000
DEFINITION  uy59b09.y1 McCarrey Eddy round spermatid Mus musculus cDNA clone
            IMAGE:3663833 5' similar to TR:O54918 O54918 BCL2 INTERACTING
            MEDIATOR OF CELL DEATH ;, mRNA sequence.
ACCESSION   BF021882
VERSION     BF021882.1  GI:10753214
KEYWORDS    EST.
SOURCE      house mouse.
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 452)
  AUTHORS   Marra,M., Hillier,L., Kucaba,T., Martin,J., Beck,C., Wylie,T.,
            Underwood,K., Steptoe,M., Theising,B., Allen,M., Bowers,Y., Person
            ,B., Swaller,T., Gibbons,M., Pape,D., Harvey,N., Schurk,R., Ritter
            ,E., Kohn,S., Shin,T., Jackson,Y., Cardenas,M., McCann,R.,
            Waterston,R. and Wilson,R.
  TITLE     The WashU-NCI Mouse EST Project 1999
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Marra M/WashU-NCI Mouse EST Project 1999
            Washington University School of Medicine
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: mouseest@watson.wustl.edu
            This clone is available royalty-free through LLNL ; contact the
            IMAGE Consortium (info@image.llnl.gov) for further information.
            MGI:1424601
            Seq primer: Primer name ambiguous
            High quality sequence stop: 386.
FEATURES             Location/Qualifiers
     source          1. .452
                     /organism="Mus musculus"
                     /strain="CD-1"
                     /db_xref="taxon:10090"
                     /clone="IMAGE:3663833"
                     /clone_lib="McCarrey Eddy round spermatid"
                     /sex="male"
                     /tissue_type="round spermatids, pooled from multiple mice"
                     /dev_stage="60 day"
                     /lab_host="DH10B (phage-resistant)"
                     /note="Organ: testis; Vector: pBluescript SK+ (Stratagene
                     ); Site_1: XhoII; Site_2: EcoRI; cDNA oligo dT-primed
                     [5'-(GA)10-ACTAGTCTCGAGTTTTTTTTTTTTT-3'] and directionally
                     cloned  using 5' linkers 5'-AATTCGGCACGAG-3' and
                     5'-CTCGTGCCG-3'. Size selection of >400bp material gives
                     average insert size ranging from 1-2 kb. Library was mass
                     excised (from  lambda-UniZAP-XR) and resulting
                     single-stranded phagemids were prepped and tranformed
                     into DH10B. Library contains 98.5% recombinants.
                     References: J. Androl. 20:635-639 and Gene 25:263-269.
                     Library constructed and donated by J. McCarrey, Ph.D.
                     (Southwest Foundation for Biomedical Research, Dept. of
                     Genetics); excision done by E.M. Eddy, Ph.D. (National
                     Institutes of  Health, National Institute of Environmental
                     Health Sciences).  Original lambda-based library is
                     available through ATCC, catalog #63423."
BASE COUNT      106 a    130 c    112 g    104 t
ORIGIN
```

alignment_scores:

```
            Quality:  357.00            Length:     69
              Ratio:  5.174               Gaps:      0
 Percent Similarity: 100.000   Percent Identity: 100.000
```

alignment_block:

```
 US-09-508-832-2 x BF021882   ..
```

Align seg 1/1 to: BF021882 from: 1 to: 452

```
       42 AlaSerIleArgGlnSerGlnGluGluProGluAspLeuArgProGluIl 58
          ||||||||||||||||||||||||||||||||||||||||||||||||||
      211 GCTTCCATACGACAGTCTCAGGAGGAACCTGAAGATCTGCGCCCGGAGAT 260

       58 eArgIleAlaGlnGluLeuArgArgIleGlyAspGluPheAsnGluThrT 75
          ||||||||||||||||||||||||||||||||||||||||||||||||||
      261 ACGGATTGCACAGGAGCTGCGGCGGATCGGAGACGAGTTCAACGAAACTT 310

       75 yrThrArgArgValPheAlaAsnAspTyrArgGluAlaGluAspHisPro 91
          ||||||||||||||||||||||||||||||||||||||||||||||||||
      311 ACACAAGGAGGGTGTTTGCAAATGATTACCGCGAGGCTGAAGACCACCCT 360

       92 GlnMetValIleLeuGlnLeuLeuArgPheIlePheArgLeuValTrpAr 108
          ||||||||||||||||||||||||||||||||||||||||||||||||||
      361 CAAATGGTTATCTTACAACTGTTACGCTTTATCTTCCGTCTGGTATGGAG 410

      108 gArgHis 110
          |||||||
      411 AAGGCAT 417
```

seq_name: gb_est2:BF319454

seq_documentation_block:

```
LOCUS       BF319454      389 bp    mRNA            EST       29-DEC-2000
DEFINITION  uy59b09.x1 McCarrey Eddy round spermatid Mus musculus cDNA clone
```

```
            IMAGE:3663833 3' similar to TR:O54918 O54918 BCL2 INTERACTING
            MEDIATOR OF CELL DEATH ;, mRNA sequence.
ACCESSION   BF319454
VERSION     BF319454.1  GI:11268195
KEYWORDS    EST.
SOURCE      house mouse.
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 389)
  AUTHORS   Marra,M., Hillier,L., Kucaba,T., Martin,J., Beck,C., Wylie,T.,
            Underwood,K., Steptoe,M., Theising,B., Allen,M., Bowers,Y., Person
            ,B., Swaller,T., Gibbons,M., Pape,D., Harvey,N., Schurk,R., Ritter
            ,E., Kohn,S., Shin,T., Jackson,Y., Cardenas,M., McCann,R.,
            Waterston,R. and Wilson,R.
  TITLE     The WashU-NCI Mouse EST Project 1999
  JOURNAL   Unpublished (1999)
COMMENT     Other_ESTs: uy59b09.y1
            Contact: Marra M/WashU-NCI Mouse EST Project 1999
            Washington University School of Medicine
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: mouseest@watson.wustl.edu
            This clone is available royalty-free through LLNL ; contact the
            IMAGE Consortium (info@image.llnl.gov) for further information.
            MGI:1424601
            High quality sequence stop: 325.
FEATURES             Location/Qualifiers
     source          1. .389
                     /organism="Mus musculus"
                     /strain="CD-1"
                     /db_xref="taxon:10090"
                     /clone="IMAGE:3663833"
                     /clone_lib="McCarrey Eddy round spermatid"
                     /sex="male"
                     /tissue_type="round spermatids, pooled from multiple mice"
                     /dev_stage="60 day"
                     /lab_host="DH10B (phage-resistant)"
                     /note="Organ: testis; Vector: pBluescript SK+ (Stratagene
                     ); Site_1: XhoII; Site_2: EcoRI; cDNA oligo dT-primed
                     [5'-(GA)10-ACTAGTCTCGAGTTTTTTTTTTTT-3'] and directionally
                     cloned  using 5' linkers 5'-AATTCGGCACGAG-3' and
                     5'-CTCGTGCCG-3'. Size selection of >400bp material gives
                     average insert size ranging from 1-2 kb. Library was mass
                     excised (from  lambda-UniZAP-XR) and resulting
                     single-stranded phagemids were prepped and tranformed
                     into DH10B. Library contains 98.5% recombinants.
                     References: J. Androl. 20:635-639 and Gene 25:263-269.
                     Library constructed and donated by J. McCarrey, Ph.D.
                     (Southwest Foundation for Biomedical Research, Dept. of
                     Genetics); excision done by E.M. Eddy, Ph.D. (National
                     Institutes of  Health, National Institute of Environmental
                     Health Sciences).  Original lambda-based library is
                     available through ATCC, catalog #63423."
BASE COUNT      100 a    104 c     89 g     96 t
ORIGIN


alignment_scores:
            Quality:  249.00             Length:     53
              Ratio:  4.882                Gaps:      1
 Percent Similarity:  96.226   Percent Identity: 94.340

alignment_block:
 US-09-508-832-2 x BF319454/rev   ..

 Align seg 1/1  to reverse of: BF319454  from: 1  to: 389

      59 ArgIleAlaGlnGluLeuArgArgIleGly.AspGluPheAsnGluThrT 75
         |||:::|||||||||||||||||   |||||| ||||||||||||||||||
     389 CGATTTGCACAGGAACTTCGGGGGATCGGAAGACGAGTTCAACGAAACTT 340
```

```
      75 yrThrArgArgValPheAlaAsnAspTyrArgGluAlaGluAspHisPro 91
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     339 ACACAAGGAGGGTGTTTGCAAATGATTACCGCGAGGCTGAAGACCACCCT 290

      92 GlnMetValIleLeuGlnLeuLeuArgPheIlePheArgLeuValTrpAr 108
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     289 CAAATGGTTATCTTACAACTGTTACGCTTTATCTTCCGTCTGGTATGGAG 240

     109 gArgHis 110
         |||||||
     239 AAGGCAT 233

seq_name: gb_gss:AZ706148

seq_documentation_block:
LOCUS       AZ706148      580 bp    DNA             GSS       24-JAN-2001
DEFINITION  RPCI-23-227P3.TV RPCI-23 Mus musculus genomic clone RPCI-23-227P3,
            DNA sequence.
ACCESSION   AZ706148
VERSION     AZ706148.1  GI:12433319
KEYWORDS    GSS.
SOURCE      house mouse.
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 580)
  AUTHORS   Zhao,S., Nierman,W., Feldblyum,T., Malek,J., Shatsman,S., Akinret
            ,B., Levins,M., Mcgann,S., Tsegaye,G., Geer,K., Krol,M., de Jong,P.
            and Fraser,C.M.
  TITLE     Mouse BAC End Sequences from Library RPCI-23
  JOURNAL   Unpublished (1999)
COMMENT     Other_GSSs: RPCI-23-227P3.TJ
            Contact: Shaying Zhao
            Department of Eukaryotic Genomics
            The Institute for Genomic Research
            9712 Medical Center Dr., Rockville, MD 20850, USA
            Tel: 301 838 0200
            Fax: 301 838 0208
            Email: szhao@tigr.org
            Clones are derived from the mouse BAC library RPCI-23. For BAC
            library availability, please contact Pieter de Jong
            (pdejong@mail.cho.org). Clones may be purchased from BACPAC
            Resources (http://www.chori.org/bacpac/orderingframe.htm). BAC end
            page: http://ww .tigr.org/tdb/bac_ends/mouse/bac_end_intro.html
            Plate: 227  row: P  column: 3
            Seq primer: T7
            Class: BAC ends.
FEATURES             Location/Qualifiers
     source          1. .580
                     /organism="Mus musculus"
                     /strain="C57BL/6J"
                     /db_xref="taxon:10090"
                     /clone="RPCI-23-227P3"
                     /clone_lib="RPCI-23"
                     /sex="Female"
                     /lab_host="DH10B"
                     /note="Organ: Kidney/Brain; Vector: pBACe3.6; Site_1:
                     EcoRI; Site_2: EcoRI; Female C57BL/6J mouse kidney and/or
                     brain genomic DNA was isolated and  partially digested
                     with a combination of EcoRI and EcoRI Methylase.   Size
                     selected DNA was cloned into the pBACe3.6 vector at the
                     EcoRI  sites.  The ligation products were transformed into
                     DH10B  electrocompetent cells (BRL Life Technologies).   "
BASE COUNT      138 a    162 c    138 g    142 t
ORIGIN


alignment_scores:
            Quality:  218.00             Length:     43
              Ratio:  5.070                Gaps:      0
 Percent Similarity: 100.000   Percent Identity: 95.349

alignment_block:
```

```
 US-09-508-832-2 x AZ706148   ..

 Align seg 1/1  to: AZ706148  from: 1  to: 580

       1 MetAlaLysGlnProSerAspValSerSerGluCysAspArgGluGlyGl 17
         ||||||||||||||||||||||||||||||||||||||||||||||||||
      90 ATGGCCAAGCAACCTTCTGATGTAAGTTCTGAGTGTGACAGAGAAGGTGG 139

      17 yGlnLeuGlnProAlaGluArgProProGlnLeuArgProGlyAlaProT 34
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     140 ACAATTGCAGCCTGCTGAGAGGCCTCCCCAGCTCAGGCCTGGGGCCCCTA 189

      34 hrSerLeuGlnThrGluProGlnAlaSer 43
         ||||||||||||||||||||||::::::
     190 CCTCCCTACAGACAGAACCGCAAGGTAAT 218

seq_name: gb_est2:BG173095

seq_documentation_block:
LOCUS       BG173095      668 bp    mRNA            EST       06-FEB-2001
DEFINITION  602336666F1 NCI_CGAP_Mam1 Mus musculus cDNA clone IMAGE:4459720 5',
            mRNA sequence.
ACCESSION   BG173095
VERSION     BG173095.1  GI:12679707
KEYWORDS    EST.
SOURCE      house mouse.
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 668)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Gilbert Smith, Ph.D.
            cDNA Library Preparation: Life Technologies, Inc.
            cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
            DNA Sequencing by: Incyte Genomics, Inc.
            Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLAM10260  row: c  column: 17
            High quality sequence stop: 599.
FEATURES             Location/Qualifiers
     source          1. .668
                     /organism="Mus musculus"
                     /strain="FVB/N"
                     /db_xref="taxon:10090"
                     /clone="IMAGE:4459720"
                     /clone_lib="NCI_CGAP_Mam1"
                     /tissue_type="tumor, biopsy sample"
                     /dev_stage="3 months, virgin"
                     /lab_host="DH10B"
                     /note="Organ: mammary; Vector: pCMV-SPORT6; Site_1: SalI;
                     Site_2: NotI; Cloned unidirectionally.  Primer: Oligo dT.
                     Library constructed by Life Technologies. Investigator
                     providing samples: Gilbert Smith, NIH"
BASE COUNT      135 a    235 c    162 g    136 t
ORIGIN


alignment_scores:
            Quality:  218.00            Length:     43
              Ratio:  5.070               Gaps:      0
 Percent Similarity: 100.000   Percent Identity: 95.349

alignment_block:
 US-09-508-832-2 x BG173095   ..

 Align seg 1/1  to: BG173095  from: 1  to: 668

       1 MetAlaLysGlnProSerAspValSerSerGluCysAspArgGluGlyGl 17
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     223 ATGGCCAAGCAACCTTCTGATGTAAGTTCTGAGTGTGACAGAGAAGGTGG 272

      17 yGlnLeuGlnProAlaGluArgProProGlnLeuArgProGlyAlaProT 34
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     273 ACAATTGCAGCCTGCTGAGAGGCCTCCCCAGCTCAGGCCTGGGGCCCCTA 322

      34 hrSerLeuGlnThrGluProGlnAlaSer 43
         ||||||||||||||||||||||::::::
     323 CCTCCCTACAGACAGAACCGCAAGGTAAT 351

seq_name: gb_est1:AI971169

seq_documentation_block:
LOCUS       AI971169      492 bp    mRNA            EST       08-MAR-2000
DEFINITION  wr24h12.x1 NCI_CGAP_Pr28 Homo sapiens cDNA clone IMAGE:2488679 3'
            similar to TR:O43522 O43522 BIML. [1] ;, mRNA sequence.
ACCESSION   AI971169
VERSION     AI971169.1  GI:5767995
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 492)
  AUTHORS   NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
  TITLE     National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
            Tumor Gene Index
  JOURNAL   Unpublished (1997)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Michael J. Brownstein, M.D., Ph.D., Michael R.
            Emmert-Buck, M.D., Ph.D.
            cDNA Library Preparation: M. Bento Soares, Ph.D.
            cDNA Library Arrayed by: Greg Lennon, Ph.D.
            DNA Sequencing by: Washington University Genome Sequencing Center
            Clone distribution: NCI-CGAP clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            www-bio.llnl.gov/bbrp/image/image.html
            Insert Length: 712   Std Error: 0.00
            Seq primer: -40UP from Gibco
            High quality sequence stop: 450.
FEATURES             Location/Qualifiers
     source          1. .492
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:2488679"
                     /clone_lib="NCI_CGAP_Pr28"
                     /sex="male"
                     /dev_stage="adult"
                     /lab_host="DH10B"
                     /note="Organ: prostate; Vector: pT7T3D-Pac (Pharmacia)
                     with a modified polylinker; Plasmid DNA from the
                     normalized library NCI_CGAP_Pr22 was prepared, and ss
                     circles were made in vitro. Following HAP purification,
                     this DNA was used as tracer in a subtractive hybridization
                     reaction. The driver was PCR-amplified cDNAs from a pool
                     of 5,000 clones made from the same library (cloneIDs
                     985608-986759, 1101192-1101959, and 1217928-1220615).
                     Subtraction by Bento Soares and M. Fatima Bonaldo.  "
BASE COUNT      119 a    107 c    130 g    134 t      2 others
ORIGIN


alignment_scores:
            Quality:  209.00            Length:     41
              Ratio:  5.225               Gaps:      0
 Percent Similarity:  97.561   Percent Identity: 97.561

alignment_block:
 US-09-508-832-2 x AI971169/rev   ..

 Align seg 1/1  to reverse of: AI971169  from: 1  to: 492
```

```
      1 MetAlaLysGlnProSerAspValSerSerGluCysAspArgGluGlyGl 17
        ||||||||||||||||||||||||||||||||||||||||||||||||||
    425 ATGGCAAAGCAACCTTCTGATGTAAGTTCTGAGTGTGACCGAGAAGGTAG 376

     17 yGlnLeuGlnProAlaGluArgProProGlnLeuArgProGlyAlaProT 34
        ||||||||||||||||||||||||||||||||||||||||||||||||||
    375 ACAATTGCAGCCTGCGGAGAGGCCTCCCCAGCTCAGACCTGGGGCCCCTA 326

     34 hrSerLeuGlnThrGluProGln 41
        |||||||||||||||||||||||
    325 CCTCCCTACAGACAGAGCCACAA 303

seq_name: gb_est1:AA629308

seq_documentation_block:
LOCUS       AA629308      501 bp    mRNA            EST       16-OCT-1997
DEFINITION  zu84g06.s1 Soares_testis_NHT Homo sapiens cDNA clone IMAGE:744730
            3', mRNA sequence.
ACCESSION   AA629308
VERSION     AA629308.1  GI:2541695
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 501)
  AUTHORS   Hillier,L., Allen,M., Bowles,L., Dubuque,T., Geisel,G., Jost,S.,
            Krizman,D., Kucaba,T., Lacy,M., Le,N., Lennon,G., Marra,M., Martin
            ,J., Moore,B., Schellenberg,K., Steptoe,M., Tan,F., Theising,B.,
            White,Y., Wylie,T., Waterston,R. and Wilson,R.
  TITLE     WashU-NCI human EST Project
  JOURNAL   Unpublished (1997)
COMMENT     Contact: Wilson RK
            Washington University School of Medicine
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: est@watson.wustl.edu
            This clone is available royalty-free through LLNL ; contact the
            IMAGE Consortium (info@image.llnl.gov) for further information.
            Seq primer: -40m13 fwd. ET from Amersham
            High quality sequence stop: 471.
FEATURES             Location/Qualifiers
     source          1. .501
                     /organism="Homo sapiens"
                     /db_xref="GDB:5932418"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:744730"
                     /clone_lib="Soares_testis_NHT"
                     /sex="male"
                     /lab_host="DH10B"
                     /note="Vector: pT7T3D-Pac (Pharmacia) with a modified
                     polylinker; Site_1: Not I; Site_2: Eco RI; 1st strand cDNA
                     was prepared from mRNA obtained from Clontech Laboratories
                     , Inc., and primed with  a Not I - oligo(dT) primer [5'
                     TGTTACCAATCTGAAGTGGGAGCGGCCGCCCAATTTTTTTTTTTTTTTT 3']. 
                     Double-stranded cDNA was ligated to Eco RI adaptors
                     (Pharmacia), digested with Not I and cloned into the Not I
                     and Eco RI sites of the modified pT7T3 vector. Library
                     went through one round of normalization to Cot5, and was
                     constructed  by Bento Soares and M. Fatima Bonaldo. "
BASE COUNT      155 a    112 c     97 g    137 t
ORIGIN


alignment_scores:
            Quality:  127.00             Length:     32
              Ratio:  4.536                Gaps:      0
 Percent Similarity:  87.500   Percent Identity: 71.875

alignment_block:
 US-09-508-832-2 x AA629308/rev   ..
```

```
 Align_seq 1/1  to reverse of: AA629308  from: 1  to: 501

     79 ValPheAlaAsnAspTyrArgGluAlaGluAspHisProGlnMetValIl 95
        ||||||  |||:::|||:::   ||||||||||||||||:::|||||||||
    501 GTATTTTTGAATAATTACCAAGCAGCCGAAGACCACCCACGAATGGTTAT 452

     95 eLeuGlnLeuLeuArgPheIlePheArgLeuValTrpArgArgHis 110
        .||||:::|||||||||:::|||   |||||||||||||||   |||
    451 CTTACGACTGTTACGTTACATTGTCCGCCTGGTGTGGAGAATGCAT 406

seq_name: gb_est2:BF172831

seq_documentation_block:
LOCUS       BF172831      210 bp    mRNA            EST       23-MAR-2001
DEFINITION  PCL5805 Myeloma (PCL) cDNA library Homo sapiens cDNA, mRNA
            sequence.
ACCESSION   BF172831
VERSION     BF172831.1  GI:13439045
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 210)
  AUTHORS   Claudio,J.O., Tang,H., Khan,E.M., Voralia,M., Li,Z., Cukerman,E.,
            Franciso-Pabalan,O., Liew,C.C. and Stewart,A.K.
  TITLE     The transcriptional phenotype of myeloma cells
  JOURNAL   Unpublished (2000)
COMMENT     Contact: A. Keith Stewart, M.D.
            Oncology Research
            University Health Network
            610 University Ave., 5-126, Toronto, Ontario, M5G 2M9, Canada
            Tel: (416) 946-4639
            Fax: (416) 946-6546
            Email: k.stewart@utoronto.ca
            PCR PRimers
            FORWARD: 5'-GCCAAGCTCGAAATTAACCCTCACTAAAGGG-3'
            BACKWARD: 5'-CCAGTGAATTGTAATACGACTCACTATAGGGCG-3'
            Seq primer: 5'-GAAATTAACCCTCACTAAAGG-3'.
FEATURES             Location/Qualifiers
     source          1. .210
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone_lib="Myeloma (PCL) cDNA library"
                     /sex="male"
                     /tissue_type="Blood"
                     /cell_type="myeloma"
                     /dev_stage="Plasma cell leukemia"
                     /note="Vector: Lambda Zap Express; Site_1: EcoRI; Site_2:
                     XhoI; mRNA was purified from plasma cell leukemia
                     patient's peripheral blood  containing >95% myeloma. An
                     oligo d(T)18 primer containing XhoI restriction  site was
                     used to prime first strand synthesis using M-MLV reverse
                     transcriptase. To protect the cDNAs from XhoI digestion in
                     subsequent  cloning step, the nucleotide analogue
                     5-methyl-dCTP was added to the  nucleotide mixture and
                     [a-32P]dATP was added to monitor the quantity and  quality
                     of first strand synthesis. After second-strand synthesis
                     and  blunting of cDNA termini, EcoRI adapters were ligated
                     , followed by kinase  treatment and digestion with XhoI.
                     The cDNAs were then size-fractionated  using Sephacryl
                     S-500 column and then ligated into EcoRI and XhoI digested
                     Lambda Zap Express vector. The ligation product was
                     packaged using Gigapack  II packaging extract. The library
                     had primary titre of approx. 1x106. Clones from  the
                     primary library were randomly selected for single pass
                     sequencing."
BASE COUNT      49 a     40 c     57 g     59 t      5 others
ORIGIN


alignment_scores:
```

```
           Quality:  126.00            Length:     39
             Ratio:  3.818               Gaps:      2
 Percent Similarity:  84.615   Percent Identity: 76.923

alignment_block:
 US-09-508-832-2 x BF172831   ..

 Align seg 1/1  to: BF172831  from: 1  to: 210

      42 AlaSerIleArgGlnSerGlnGluGluProGluAspLeuArgProGluIl 58
         ||||||:::|||||||||      ||||||   |||:::||||||||||
      95 GCTTCCATGAGGCAGTCT......GAACCTGCAGATATGCGCCCAGAGAT 138

      58 eArgIleAlaGlnGluLeu.ArgArgIleGlyAspGluPheAsnGluThr 74
         |   ||||||||||||||| |||||||||||||||||||||||   :::
     139 ATGGATCGCCCAAGAGTTGGCGGCGTATTGGAGACGAGTTTAACGCTAAC 188

      75 TyrThrArgArgVal 79
         ||||||||||||||||
     189 TATACAAGGAGGGTA 203

seq_name: gb_est1:AA629050

seq_documentation_block:
LOCUS       AA629050      501 bp    mRNA            EST       16-OCT-1997
DEFINITION  zu84a06.s1 Soares_testis_NHT Homo sapiens cDNA clone IMAGE:744658
            3', mRNA sequence.
ACCESSION   AA629050
VERSION     AA629050.1  GI:2541437
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 501)
  AUTHORS   Hillier,L., Allen,M., Bowles,L., Dubuque,T., Geisel,G., Jost,S.,
            Krizman,D., Kucaba,T., Lacy,M., Le,N., Lennon,G., Marra,M., Martin
            ,J., Moore,B., Schellenberg,K., Steptoe,M., Tan,F., Theising,B.,
            White,Y., Wylie,T., Waterston,R. and Wilson,R.
  TITLE     WashU-NCI human EST Project
  JOURNAL   Unpublished (1997)
COMMENT     Contact: Wilson RK
            Washington University School of Medicine
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: est@watson.wustl.edu
            This clone is available royalty-free through LLNL ; contact the
            IMAGE Consortium (info@image.llnl.gov) for further information.
            Seq primer: -40m13 fwd. ET from Amersham
            High quality sequence stop: 474.
FEATURES             Location/Qualifiers
     source          1. .501
                     /organism="Homo sapiens"
                     /db_xref="GDB:5932318"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:744658"
                     /clone_lib="Soares_testis_NHT"
                     /sex="male"
                     /lab_host="DH10B"
                     /note="Vector: pT7T3D-Pac (Pharmacia) with a modified
                     polylinker; Site_1: Not I; Site_2: Eco RI; 1st strand cDNA
                     was prepared from mRNA obtained from Clontech Laboratories
                     , Inc., and primed with  a Not I - oligo(dT) primer [5'
                     TGTTACCAATCTGAAGTGGGAGCGGCCGCCCAATTTTTTTTTTTTTTTT 3'].
                     Double-stranded cDNA was ligated to Eco RI adaptors
                     (Pharmacia), digested with Not I and cloned into the Not I
                     and Eco RI sites of the modified pT7T3 vector. Library
                     went through one round of normalization to Cot5, and was
                     constructed  by Bento Soares and M. Fatima Bonaldo. "
BASE COUNT      155 a    111 c     98 g    137 t
ORIGIN

alignment_scores:
           Quality:  123.00            Length:     31
             Ratio:  4.556               Gaps:      0
 Percent Similarity:  87.097   Percent Identity: 70.968

alignment_block:
 US-09-508-832-2 x AA629050/rev   ..

 Align seg 1/1  to reverse of: AA629050  from: 1  to: 501

      80 PheAlaAsnAspTyrArgGluAlaGluAspHisProGlnMetValIleLe 96
         |||   |||:::|||:::   |||||||||||||||:::||||||||||
     499 TTTTTGAATAATTACCAAGCAGCCGAAGACCACCCACGAATGGTTATCTT 450

      96 uGlnLeuLeuArgPheIlePheArgLeuValTrpArgArgHis 110
         |:::||||||||:::|||   |||||||||||||||   |||
     449 ACGACTGTTACGTTACATTGTCCGCCTGGTGTGGAGAATGCAT 407

seq_name: gb_est1:AW629314

seq_documentation_block:
LOCUS       AW629314      664 bp    mRNA            EST       31-MAR-2000
DEFINITION  hi56e02.x1 Soares_NFL_T_GBC_S1 Homo sapiens cDNA clone
            IMAGE:2976314 3' similar to TR:O43522 O43522 BIML. [1] ;, mRNA
            sequence.
ACCESSION   AW629314
VERSION     AW629314.1  GI:7376104
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 664)
  AUTHORS   NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
  TITLE     National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
            Tumor Gene Index
  JOURNAL   Unpublished (1997)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            This clone is available royalty-free through LLNL ; contact the
            IMAGE Consortium (info@image.llnl.gov) for further information.
            Seq primer: -40UP from Gibco
            High quality sequence stop: 458.
FEATURES             Location/Qualifiers
     source          1. .664
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:2976314"
                     /clone_lib="Soares_NFL_T_GBC_S1"
                     /lab_host="DH10B"
                     /note="Organ: pooled; Vector: pT7T3D-Pac (Pharmacia) with
                     a modified polylinker; Site_1: Not I; Site_2: Eco RI;
                     Equal amounts of plasmid DNA from three normalized
                     libraries (fetal lung NbHL19W, testis NHT, and B-cell
                     NCI_CGAP_GCB1) were mixed, and ss circles were made in
                     vitro. Following HAP purification, this DNA was used as
                     tracer in a subtractive hybridization reaction. The driver
                     was PCR-amplified cDNAs from pools of 5,000 clones made
                     from the same 3 libraries. The pools consisted of
                     I.M.A.G.E. clones 297480-302087, 682632-687239,
                     726408-728711, and 729096-731399. Subtraction by Bento
                     Soares and M. Fatima Bonaldo.  "
BASE COUNT      176 a    131 c    148 g    208 t      1 others
ORIGIN

alignment_scores:
           Quality:  123.00            Length:     39
             Ratio:  4.241               Gaps:      0
 Percent Similarity:  74.359   Percent Identity: 64.103

alignment_block:
```

```
 US-09-508-832-2 x AW629314/rev   ..

 Align seg 1/1  to reverse of: AW629314  from: 1  to: 664

       3 LysGlnProSerAspValSerSerGluCysAspArgGluGlyGlyGlnLe 19
         |||   |||   |||   |||   ||||||      ||||||   |||::
     664 AAGCCACCTTTTGATGAAAGTTTTGAGTGTGCCNGAGAAGGTAGACAAAT 615

      19 uGlnProAlaGluArgProProGlnLeuArgProGlyAlaProThrSerL 36
         :|||||||||:::|||   ||||||:::|||   |||||||||||||||
     614 GCAGCCTGCGAAGAGGCTTCCCCAGTTCAGACTTGGGGCCCCTACCTCCT 565

      36 euGlnThrGluProGln 41
            :::|||||||||||||
     564 ACAAGACAGAGCCACAA 548

seq_name: gb_est1:AF209718

seq_documentation_block:
LOCUS       AF209718      537 bp    mRNA            EST       30-MAY-2000
DEFINITION  AF209718 Xenopus laevis intestine adult Xenopus laevis cDNA clone
            pXLG10 similar to Mus musculus BimEL, mRNA sequence.
ACCESSION   AF209718
VERSION     AF209718.1  GI:8110110
KEYWORDS    EST.
SOURCE      African clawed frog.
  ORGANISM  Xenopus laevis
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidea; Pipidae;
            Xenopodinae; Xenopus.
REFERENCE   1  (bases 1 to 537)
  AUTHORS   Minter,R.
  TITLE     Development of Antibody Technology to Identify Natural Killer Cell
            Surface Antigens in Xenopus laevis. Thesis (1999) University of
            Durham, South Rd., Durham, UK
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Minter,R., Horton,J.D. and Watson,M.D.
            Biological Sciences
            University of Durham
            South Road, Durham, DH1 3LE, UK
            Email: martin.watson@durham.ac.uk.
FEATURES             Location/Qualifiers
     source          1. .537
                     /organism="Xenopus laevis"
                     /db_xref="taxon:8355"
                     /clone="pXLG10"
                     /clone_lib="Xenopus laevis intestine adult"
                     /tissue_type="intestine"
                     /cell_type="epithelial lymphocyte"
                     /dev_stage="adult"
BASE COUNT      126 a    129 c    143 g    139 t
ORIGIN


alignment_scores:
           Quality:   98.00            Length:     53
             Ratio:  2.579              Gaps:      3
 Percent Similarity:  71.698   Percent Identity: 52.830

alignment_block:
 US-09-508-832-2 x AF209718   ..

 Align seg 1/1  to: AF209718  from: 1  to: 537

       1 MetAlaLysGlnProSerAspValSerSerGluCysAspArg...GluGl 16
         ||||||||||||||||||   :::|||   ||||||:::      |||||
     223 ATGGCCAAACAACCGTCGGTCTTGAGTCCGGAGTGTAATAGTGGTGAAGG 272

      16 yGlyGlnLeuGlnProAlaGluArgProProGlnLeuArgPro..... 30
         ||||||||||||   :::::::|||      :::   ||||||
     273 TGGCCAGTTACAATCAACAAGCAGACGACATTCTCATCGTCCTCGCAGAA 322

      31 ..GlyAlaProThrSerLeuGlnThrGluProGlnAlaSerIleArgGln 46

         ||||||||||||||||||:::::      |||::::::      |||
     323 GAGGGGCCCCCACCTCTCTTAGCAGTCCTTTTCAAGGTAAT......CAA 366

      47 SerGlnGlu 49
         |||:::|||
     367 TCAGATGAG 375

seq_name: gb_est2:BG990772

seq_documentation_block:
LOCUS       BG990772      289 bp    mRNA            EST       13-JUN-2001
DEFINITION  MR1-HT1136-260101-015-a12 HT1136 Homo sapiens cDNA, mRNA sequence.
ACCESSION   BG990772
VERSION     BG990772.1  GI:14394842
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 289)
  AUTHORS   Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R.,
            Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F.,
            Goldman,G.H., Carvalho,A.F., Matsukuma,A., Baia,G.S., Simpson,D.H.,
            Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V., O'Hare
            ,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and
            Simpson,A.J.
  TITLE     Shotgun sequencing of the human transcriptome with ORF expressed
            sequence tags
  JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
  MEDLINE   20202663
COMMENT     Contact: Simpson A.J.G.
            Laboratory of Cancer Genetics
            Ludwig Institute for Cancer Research
            Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
            Brazil
            Tel: +55-11-2704922
            Fax: +55-11-2707001
            Email: asimpson@ludwig.org.br
            This sequence was derived from the FAPESP/LICR Human Cancer Genome
            Project. This entry can be seen in the following URL
            (http://www.ludwig.org.br/scripts/gethtml2.pl?t1=MR1&t2=MR1-HT1136-
            260101-015-a12&t3=2001-01-26&t4=1)
            Seq primer: puc 18 forward
            High quality sequence start: 22
            High quality sequence stop: 289.
FEATURES             Location/Qualifiers
     source          1. .289
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone_lib="HT1136"
                     /dev_stage="Adult"
                     /note="Organ: head_neck; Vector: puc18; Site_1: SmaI;
                     Site_2: SmaI; A mini-library was made by cloning products
                     derived from  ORESTES PCR (U.S. Letters Patent application
                     No. 196,716 -  Ludwig Institute for Cancer Research)
                     profiles into  the pUC 18 vector. Reverse transcription of
                     tissue mRNA and cDNA amplification were performed under
                     low stringency conditions."
BASE COUNT       84 a     55 c     78 g     72 t
ORIGIN


alignment_scores:
           Quality:   92.00            Length:     23
             Ratio:  4.381              Gaps:      0
 Percent Similarity:  91.304   Percent Identity: 69.565

alignment_block:
 US-09-508-832-2 x BG990772/rev   ..

 Align seg 1/1  to reverse of: BG990772  from: 1  to: 289

      88 GluAspHisProGlnMetValIleLeuGlnLeuLeuArgPheIlePheAr 104
```

```
      ::::::||||||:::||||||||||||:::||||||||:::|||   ||
    170 CAAAACCACCCACGAATGGTTATCTTACGACTGTTACGTTACATTGTCCG 121

    104 gLeuValTrpArgArgHis 110
        ||||||||||||   |||
    120 CCTGGTGTGGAGAATGCAT 102

seq_name: gb_gss:CNS0085G

seq_documentation_block:
LOCUS       CNS0085G      862 bp    DNA             GSS       03-JUN-1999
DEFINITION  Drosophila melanogaster genome survey sequence TET3 end of BAC #
            BACR16K05 of RPCI-98 library from Drosophila melanogaster (fruit
            fly), genomic survey sequence.
ACCESSION   AL051751
VERSION     AL051751.1  GI:4933300
KEYWORDS    GSS.
SOURCE      fruit fly.
  ORGANISM  Drosophila melanogaster
            Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
            Pterygota; Neoptera; Endopterygota; Diptera; Brachycera;
            Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.
REFERENCE   1  (bases 1 to 862)
  AUTHORS   Genoscope.
  TITLE     Direct Submission
  JOURNAL   Submitted (02-JUN-1999) Genoscope - Centre National de Sequencage :
            BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr
            - Web : www.genoscope.cns.fr)
COMMENT     Determination of this BAC-end sequence was carried out as part of a
            collaboration with the Berkeley Drosophila Genome Project (BDGP).
            The BDGP is constructing a physical map of the Drosophila
            melanogaster genome using these BACs. For further information
            please see  http://www.fruitfly.org The BDGP Drosophila
            melanogaster BAC library was prepared by Kazutoyo Osoegawa and
            Aaron Mammoser in Pieter de Jong's laboratory in the Department of
            Cancer Genetics at the Roswell Park Cancer Institute in Buffalo,
            NY. The library is named RPCI-98 and was constructed by partial
            EcoRI digestion of Drosophila DNA provided by the BDGP from the
            isogenic strain y2; cn bw sp, the same strain used for the BDGP's
            P1 and EST libraries. A more detailed description of the library
            and how to order individual BAC clones, the entire library, or
            filters for hybridization from the BACPAC Resource Center can be
            found at http://bacpac.med.buffalo.edu/drosophila_bac.htm.
FEATURES             Location/Qualifiers
     source          1. .862
                     /organism="Drosophila melanogaster"
                     /db_xref="taxon:7227"
                     /clone_lib="RPCI-98"
                     /clone="BACR16K05"
                     /note="end : TET3"
BASE COUNT      278 a    221 c    217 g     73 t     73 others
ORIGIN


alignment_scores:
           Quality:   91.00            Length:     97
             Ratio:  1.784              Gaps:      2
 Percent Similarity:  52.577   Percent Identity: 27.835

alignment_block:
 US-09-508-832-2 x CNS0085G   ..

 Align seg 1/1  to: CNS0085G  from: 1  to: 862

     14 ArgGluGlyGlyGlnLeuGlnProAlaGluArgProProGlnLeuArgPr 30
        |||:::|||||              ||||||||||||:::   |||
    215 CGCCAGGGGGGA.............GAAAGGCCCCCMCGAAGCAGAGA 249

     30 oGlyAlaProThrSerLeuGlnThrGluProGlnAlaSerIleArgGlnS 47
         |||        :::   ||||||:::   :::   :::   :::::::
    250 AGGGCCCAAGCCCAACCAACAGACGAAACAGCGACACAAAAAAAAAAACG 299

     47 erGlnGluGluProGluAsp........................... 53
```

```
      :::::|||:::||||||:::
    300 GCGACGAAAACCCAGAAGAGAAAAAGCAAAAAAAACAAACCAACAGGCCA 349

     54 ...LeuArgProGluIleArgIleAlaGlnGluLeuArgArgIleGlyAs 69
           :::|||:::   :::      :::      ||||||   :::||
    350 GCGAGAAAACCCCAAAACAAAAAACGGAAAACGCAGCGGAGGACGGCGGA 399

     69 pGluPheAsnGluThrTyrThrArgArgValPheAlaAsnAspTyrArgG 86
        ||||   |||           |||:::::      ::::::   |||:
    400 CGAGGCCAACGCAGAGGAAAAAAGAGAAACAAACGAAAAAGAACGCCGTA 449

     86 luAlaGluAspHisProGlnMetValIleLeuGlnLeuLeu 99
        ::|||:::      |||      :::   ||||||   |||
    450 GAGCACATACGGGCCCYACGGAGATTCGCCTCCAACHTTTA 490

seq_name: gb_est2:BG420514

seq_documentation_block:
LOCUS       BG420514      921 bp    mRNA            EST       14-MAR-2001
DEFINITION  602452566F1 NIH_MGC_14 Homo sapiens cDNA clone IMAGE:4590688 5',
            mRNA sequence.
ACCESSION   BG420514
VERSION     BG420514.1  GI:13327020
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 921)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: DCTD/DTP
             cDNA Library Preparation: Ling Hong/Rubin Laboratory
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Incyte Genomics, Inc.
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLCM1326  row: d  column: 17
            High quality sequence stop: 689.
FEATURES             Location/Qualifiers
     source          1. .921
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:4590688"
                     /clone_lib="NIH_MGC_14"
                     /tissue_type="renal cell adenocarcinoma"
                     /lab_host="DH10B (phage-resistant)"
                     /note="Organ: kidney; Vector: pOTB7; Site_1: XhoI; Site_2:
                     EcoRI; cDNA made by oligo-dT priming.  Directionally
                     cloned into EcoRI/XhoI sites using the following 5'
                     adaptor: GGCACGAG(G).  Size-selected >500bp for average
                     insert size 1.8kb.  Library constructed by Ling Hong in
                     the laboratory of Gerald M. Rubin (University of
                     California, Berkeley) using  ZAP-cDNA synthesis kit
                     (Stratagene) and Superscript II RT (Life Technologies)."
BASE COUNT      215 a    283 c    285 g    138 t
ORIGIN


alignment_scores:
           Quality:   89.50            Length:    109
             Ratio:  1.543              Gaps:      7
 Percent Similarity:  53.211   Percent Identity: 31.193

alignment_block:
 US-09-508-832-2 x BG420514   ..

 Align seg 1/1  to: BG420514  from: 1  to: 921
```

```
 13 AspArgGluGlyGlyGlnLeuGlnProAlaGluArgProProGlnLeuAr 29
    :::|||||||||   :::   :::|||   ||||||||||||   |||
274 GAACGAGAAGGAGACCATGCAAAGCCT...GAACGACCGCCTGGCCTCTT 320

 29 gProGlyAlaPro.ThrSerLeuGlnThrGluProGlnAlaSerIleArg 45
     ||||||          ||||||:::|||||   :::
321 ACCTGGACAGAGTGAGGAGCCTGGAGACCGAGAACCGGAGGCTGGAGAGC 370

 46 GlnSerGlnGluGluProGluAspLeuArgProGluIleArg........ 59
    :::   :::|||:::   |||         |||::::::|||
371 AAAATCCGGGAGCACTTGGAGAAGAAGGGACCCCAGGTCAGAGACTGGAG 420

 60 .............IleAlaGlnGluLeuArgArgIleGlyAspGluPheA 72
                 |||   ::::::||||||
421 CCATTACTTCAAGATCATCGAGGACCTGAGG................... 451

 72 snGluThrTyrThrArgArgValPheAlaAsnAsp.............. 83
                   ::::::||||||||||:::
452 ..............GCTCAGATCTTCGCAAATAACTGTGGACAATGCCCG 487

 84 TyrArgGluAlaGluAspHis.ProGlnMetValIleLeuGlnLeuLeuA 100
    :::|||:::|||:::|||||| |||   :::::::::|||:::   |||:
488 CATCGTTCTGCAGATGACCATGCCC...GTCTTGCTGCTGATGACTTTAG 534

100 rgPheIlePheArgLeuValTrp 107
    ::      :::|||      |||
535 AGTCAAGTATGAGACAGAGCTGG 557
```

```
OM of: US-09-508-832-4 to: GenEmbl:*   out_format : pfs

Date: Dec 11, 2001  1:45 AM

About: Results were produced by the GenCore software, version 4.5,
       Copyright (c) 1993-2000 Compugen Ltd.

Command line parameters:
 -MODEL=frame+_p2n.model -DEV=xlp
 -Q=/cgn2_1/USPTO_spool/US09508832/runat_10122001_110349_29549/app_query.fasta_1.620
 -DB=GenEmbl -QFMT=fastap -SUFFIX=rge -GAPOP=12.000 -GAPEXT=4.000
 -MINMATCH=0.100 -LOOPCL=0.000 -LOOPEXT=0.000 -QGAPOP=4.500
 -QGAPEXT=0.050 -XGAPOP=10.000 -XGAPEXT=0.500 -FGAPOP=6.000
 -FGAPEXT=7.000 -YGAPOP=10.000 -YGAPEXT=0.500 -DELOP=6.000
 -DELEXT=7.000 -START=1 -MATRIX=blosum62 -TRANS=human40.cdi
 -LIST=45 -DOCALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0
 -ALIGN=15 -MODE=LOCAL -OUTFMT=pfs -NORM=ext -MINLEN=0
 -MAXLEN=2000000000 -USER=US09508832_@CGN1_1_0 -NCPU=6 -ICPU=3
 -LONGLOG -NO_XLPXY -WAIT -THREADS=1

Search information block:
Query: US-09-508-832-4
Query length: 140
Database: GenEmbl:*
Database sequences: 1472140
Database length: -341344837
Search time (sec): 2485.920000



score_list:
Sequence           Strd Orig    ZScore    EScore Len  ! Documentation  ..
gb_pat:AX031281      +  742.00 1004.53  1.1e-47   422  ! AX031281 Sequence 3 from Patent
gb_un:AX031307       +  742.00 1004.53  1.1e-47   422  ! AX031307 Sequence 3 from Patent
gb_ro:AF032460       +  742.00 1004.51  1.1e-47   423  ! AF032460 Mus musculus BimL mRNA
gb_ro:AF136927       +  727.00  984.38  1.5e-46   423  ! AF136927 Rattus norvegicus Bcl-
gb_pat:AX031283      +  704.00  951.16  1.1e-44   590  ! AX031283 Sequence 5 from Patent
gb_un:AX031309       +  704.00  951.16  1.1e-44   590  ! AX031309 Sequence 5 from Patent
gb_ro:AF032459       +  704.00  951.15  1.1e-44   591  ! AF032459 Mus musculus BimEL mRN
gb_ro:AF065433       +  689.00  931.01  1.4e-43   591  ! AF065433 Rattus norvegicus Bcl-
gb_pat:AX031285      +  629.00  852.92  3.1e-39   416  ! AX031285 Sequence 7 from Patent
gb_un:AX031311       +  629.00  852.92  3.1e-39   416  ! AX031311 Sequence 7 from Patent
gb_pr:AF032458       +  629.00  852.91  3.1e-39   417  ! AF032458 Homo sapiens BimL mRNA
gb_pat:AX031287      +  589.00  796.70  4.2e-36   596  ! AX031287 Sequence 9 from Patent
gb_un:AX031313       +  589.00  796.70  4.2e-36   596  ! AX031313 Sequence 9 from Patent
gb_pr:AF032457       +  589.00  796.69  4.2e-36   597  ! AF032457 Homo sapiens BimEL mRN
gb_pat:AX031279      +  549.00  747.10  2.4e-33   332  ! AX031279 Sequence 1 from Patent
gb_un:AX031305       +  549.00  747.10  2.4e-33   332  ! AX031305 Sequence 1 from Patent
gb_ro:AF032461       +  549.00  747.08  2.5e-33   333  ! AF032461 Mus musculus BimS mRNA
gb_ro:AF065432       +  531.00  722.91  5.4e-32   333  ! AF065432 Rattus norvegicus Bcl-
gb_ro:AF065431       +  488.00  666.35  7.7e-29   282  ! AF065431 Rattus norvegicus Bcl-
gb_htg:AC013332      -  337.00  418.31  5.0e-15   180569  ! AC013332 Homo sapiens chromo
gb_htg:AC013332      +  137.00  149.80  4.56      180569  ! AC013332 Homo sapiens chromo
gb_htg:AC046192      -  132.00  144.39  9.13      150035  ! AC046192 Homo sapiens chromo
gb_htg:AC046192      +  128.00  139.02  18.19     150035  ! AC046192 Homo sapiens chromo
gb_ba:STMWHIB12X     +  101.50  127.85  76.12     4615  ! L22864 Streptomyces aureofacie
gb_pl:AP003023       +  100.50  102.97  1.9e+03   132470  ! AP003023 Oryza sativa genomi
gb_htg:AP003245      -  100.50  101.62  2.2e+03   160468  ! AP003245 Oryza sativa chromo
gb_htg:AC024740      -  100.00  100.16  2.7e+03   179538  ! AC024740 Homo sapiens chromo
gb_htg:AC026803      -  100.00   98.38  3.3e+03   231450  ! AC026803 Homo sapiens chromo
gb_htg:AL390837      +   99.50  108.31  933.72    51090  ! AL390837 Homo sapiens chromos
gb_htg:AC092821      -   99.50  107.68  1.0e+03   55899  ! AC092821 Homo sapiens chromos
gb_htg:AL592487      -   99.50  100.52  2.5e+03   155183  ! AL592487 Homo sapiens chromo
gb_htg:AL513016      -   99.50   98.64  3.2e+03   202922  ! AL513016 Homo sapiens chromo
gb_pr:AC010616       -   99.00   99.13  3.0e+03   171849  ! AC010616 Homo sapiens chromo
gb_ro:AC019153       +   98.50   97.57  3.7e+03   195068  ! AC019153 Mus musculus chromo
gb_pr:AC008403       -   98.00   97.08  3.9e+03   190075  ! AC008403 Homo sapiens chromo
gb_in:PFAVAR23A      +   96.50  110.11  740.85    22243  ! L40609 Plasmodium falciparum,
gb_htg:AC010327      +   96.50   96.79  4.1e+03   148679  ! AC010327 Homo sapiens chromo
gb_pr:AF166330       -   95.00  113.77  463.51    9909  ! AF166330 Homo sapiens stratum
gb_htg:AC011483      +   95.00   96.94  4.0e+03   109124  ! AC011483 Homo sapiens chromo
gb_htg:AC027602      +   95.00   92.11  7.5e+03   217346  ! AC027602 Homo sapiens chromo
gb_pr:AF243527       +   95.00   91.72  7.8e+03   230000  ! AF243527 Homo sapiens serine

gb_htg:AC020922      +   94.50   94.80  5.3e+03   134693  ! AC020922 Homo sapiens chr
gb_htg:AC024486      -   94.50   93.56  6.2e+03   160737  ! AC024486 Homo sapiens chr
gb_htg:AC014673      -   94.00  105.72  1.3e+03   25789  ! AC014673 Drosophila melano
gb_in:AC008189       -   94.00   91.77  7.8e+03   188359  ! AC008189 Drosophila melan

seq_name: gb_pat:AX031281

seq_documentation_block:
LOCUS       AX031281      422 bp    DNA             PAT       20-SEP-2000
DEFINITION  Sequence 3 from Patent WO9914321.
ACCESSION   AX031281
VERSION     AX031281.1  GI:10278612
KEYWORDS    .
SOURCE      unidentified.
  ORGANISM  unidentified
            unclassified.
REFERENCE   1  (bases 1 to 422)
  AUTHORS   O'Reilly,L., Puthalakath,H., Adams,J., O'Connor,L., Cory,S.,
            Huang,D.C. and Strasser,A.
  TITLE     Novel therapeutic molecules
  JOURNAL   Patent: WO 9914321-A 3 25-MAR-1999;
            INST MEDICAL W & E HALL (AU) ; PUTHALAKATH HAMSA (AU) ; REILLY
            LORRAINE O (AU) ; ADAMS JERRY (AU) ; CONNOR LIAM O (AU) ; CORY
            SUZANNE (AU) ; HUANG DAVID C S (AU) ; STRASSER ANDREAS (AU)
FEATURES             Location/Qualifiers
     source          1. .422
                     /organism="unidentified"
                     /db_xref="taxon:32644"
     CDS             1. .>420
                     /note="unnamed protein product"
                     /codon_start=1
                     /protein_id="CAC09654.1"
                     /db_xref="GI:10278613"
                     /translation="MAKQPSDVSSECDREGGQLQPAERPPQLRPGAPTSLQTEPQDRS
                     PAPMSCDKSTQTPSPPCQAFNHYLSAMASIRQSQEEPEDLRPEIRIAQELRRIGDEFN
                     ETYTRRVFANDYREAEDHPQMVILQLLRFIFRLVWRRH"
BASE COUNT      112 a    116 c    109 g     85 t
ORIGIN


alignment_scores:
           Quality:  742.00              Length:    140
             Ratio:  5.300                 Gaps:      0
 Percent Similarity: 100.000   Percent Identity: 100.000

alignment_block:
 US-09-508-832-4 x AX031281   ..

 Align seg 1/1  to: AX031281  from: 1  to: 422

       1 MetAlaLysGlnProSerAspValSerSerGluCysAspArgGluGlyGl 17
         ||||||||||||||||||||||||||||||||||||||||||||||||||
       1 ATGGCCAAGCAACCTTCTGATGTAAGTTCTGAGTGTGACAGAGAAGGTGG 50

      17 yGlnLeuGlnProAlaGluArgProProGlnLeuArgProGlyAlaProT 34
         ||||||||||||||||||||||||||||||||||||||||||||||||||
      51 ACAATTGCAGCCTGCTGAGAGGCCTCCCCAGCTCAGGCCTGGGGCCCCTA 100

      34 hrSerLeuGlnThrGluProGlnAspArgSerProAlaProMetSerCys 50
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     101 CCTCCCTACAGACAGAACCGCAAGACAGGAGCCCGGCACCCATGAGTTGT 150

      51 AspLysSerThrGlnThrProSerProProCysGlnAlaPheAsnHisTy 67
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     151 GACAAGTCAACACAAACCCCAAGTCCTCCTTGCCAGGCCTTCAACCACTA 200

      67 rLeuSerAlaMetAlaSerIleArgGlnSerGlnGluGluProGluAspL 84
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     201 TCTCAGTGCAATGGCTTCCATACGACAGTCTCAGGAGGAACCTGAAGATC 250

      84 euArgProGluIleArgIleAlaGlnGluLeuArgArgIleGlyAspGlu 100
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     251 TGCGCCCGGAGATACGGATTGCACAGGAGCTGCGGCGGATCGGAGACGAG 300
```

```
     101 PheAsnGluThrTyrThrArgArgValPheAlaAsnAspTyrArgGluAl 117
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     301 TTCAACGAAACTTACACAAGGAGGGTGTTTGCAAATGATTACCGCGAGGC 350

     117 aGluAspHisProGlnMetValIleLeuGlnLeuLeuArgPheIlePheA 134
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     351 TGAAGACCACCCTCAAATGGTTATCTTACAACTGTTACGCTTTATCTTCC 400

     134 rgLeuValTrpArgArgHis 140
         ||||||||||||||||||
     401 GTCTGGTATGGAGAAGGCAT 420

seq_name: gb_un:AX031307

seq_documentation_block:
LOCUS       AX031307      422 bp    DNA             UNA       20-SEP-2000
DEFINITION  Sequence 3 from Patent WO9914321.
ACCESSION   AX031307
VERSION     AX031307.1  GI:10278635
KEYWORDS    .
SOURCE      unidentified.
  ORGANISM  unidentified
            unclassified.
REFERENCE   1  (bases 1 to 422)
  AUTHORS   O'Reilly,L., Puthalakath,H., Adams,J., O'Connor,L., Cory,S.,
            Huang,D.C. and Strasser,A.
  TITLE     Novel therapeutic molecules
  JOURNAL   Patent: WO 9914321-A 25-MAR-1999;
            INST MEDICAL W & E HALL (AU) ; PUTHALAKATH HAMSA (AU) ; REILLY
            LORRAINE O (AU) ; ADAMS JERRY (AU) ; CONNOR LIAM O (AU) ; CORY
            SUZANNE (AU) ; HUANG DAVID C S (AU) ; STRASSER ANDREAS (AU)
FEATURES             Location/Qualifiers
     source          1. .422
                     /organism="unidentified"
                     /db_xref="taxon:32644"
     CDS             1. .>420
                     /note="unnamed protein product"
                     /codon_start=1
                     /protein_id="CAC09659.1"
                     /db_xref="GI:10278636"
                     /translation="MAKQPSDVSSECDREGGQLQPAERPPQLRPGAPTSLQTEPQDRS
                     PAPMSCDKSTQTPSPPCQAFNHYLSAMASIRQSQEEPEDLRPEIRIAQELRRIGDEFN
                     ETYTRRVFANDYREAEDHPQMVILQLLRFIFRLVWRRH"
BASE COUNT      112 a    116 c    109 g     85 t
ORIGIN


alignment_scores:
            Quality:  742.00             Length:    140
              Ratio:  5.300                Gaps:      0
 Percent Similarity: 100.000   Percent Identity: 100.000

alignment_block:
 US-09-508-832-4 x AX031307   ..

 Align seg 1/1  to: AX031307  from: 1  to: 422

       1 MetAlaLysGlnProSerAspValSerSerGluCysAspArgGluGlyGl 17
         ||||||||||||||||||||||||||||||||||||||||||||||||||
       1 ATGGCCAAGCAACCTTCTGATGTAAGTTCTGAGTGTGACAGAGAAGGTGG 50

      17 yGlnLeuGlnProAlaGluArgProProGlnLeuArgProGlyAlaProT 34
         ||||||||||||||||||||||||||||||||||||||||||||||||||
      51 ACAATTGCAGCCTGCTGAGAGGCCTCCCCAGCTCAGGCCTGGGGCCCCTA 100

      34 hrSerLeuGlnThrGluProGlnAspArgSerProAlaProMetSerCys 50
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     101 CCTCCCTACAGACAGAACCGCAAGACAGGAGCCCGGCACCCATGAGTTGT 150

      51 AspLysSerThrGlnThrProSerProProCysGlnAlaPheAsnHisTy 67
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     151 GACAAGTCAACACAAACCCCAAGTCCTCCTTGCCAGGCCTTCAACCACTA 200

      67 rLeuSerAlaMetAlaSerIleArgGlnSerGlnGluGluProGluAspL 84
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     201 TCTCAGTGCAATGGCTTCCATACGACAGTCTCAGGAGGAACCTGAAGATC 250

      84 euArgProGluIleArgIleAlaGlnGluLeuArgArgIleGlyAspGlu 100
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     251 TGCGCCCGGAGATACGGATTGCACAGGAGCTGCGGCGGATCGGAGACGAG 300

     101 PheAsnGluThrTyrThrArgArgValPheAlaAsnAspTyrArgGluAl 117
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     301 TTCAACGAAACTTACACAAGGAGGGTGTTTGCAAATGATTACCGCGAGGC 350

     117 aGluAspHisProGlnMetValIleLeuGlnLeuLeuArgPheIlePheA 134
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     351 TGAAGACCACCCTCAAATGGTTATCTTACAACTGTTACGCTTTATCTTCC 400

     134 rgLeuValTrpArgArgHis 140
         ||||||||||||||||||
     401 GTCTGGTATGGAGAAGGCAT 420

seq_name: gb_ro:AF032460

seq_documentation_block:
LOCUS       AF032460      423 bp    mRNA            ROD       19-FEB-1998
DEFINITION  Mus musculus BimL mRNA, complete cds.
ACCESSION   AF032460
VERSION     AF032460.1  GI:2895501
KEYWORDS    .
SOURCE      house mouse.
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 423)
  AUTHORS   O'Connor,L., Strasser,A., O'Reilly,L.A., Hausmann,G., Adams,J.M.,
            Cory,S. and Huang,D.C.
  TITLE     Bim: a novel member of the Bcl-2 family that promotes apoptosis
  JOURNAL   EMBO J. 17 (2), 384-395 (1998)
  MEDLINE   98094360
   PUBMED   9430630
REFERENCE   2  (bases 1 to 423)
  AUTHORS   O'Connor,L., Strasser,A., O'Reilly,L.A., Hausmann,G., Adams,J.M.,
            Cory,S. and Huang,D.C.S.
  TITLE     Direct Submission
  JOURNAL   Submitted (03-NOV-1997) Molecular Genetics of Cancer, The Walter &
            Eliza Hall Institute of Medical Research, PO Royal Melbourne
            Hospital, Parkville, Victoria 3050, Australia
FEATURES             Location/Qualifiers
     source          1. .423
                     /organism="Mus musculus"
                     /db_xref="taxon:10090"
     CDS             1. .423
                     /note="pro-apoptotic BH3-containing Bcl-2 family member"
                     /codon_start=1
                     /product="BimL"
                     /protein_id="AAC40030.1"
                     /db_xref="GI:2895502"
                     /translation="MAKQPSDVSSECDREGGQLQPAERPPQLRPGAPTSLQTEPQDRS
                     PAPMSCDKSTQTPSPPCQAFNHYLSAMASIRQSQEEPEDLRPEIRIAQELRRIGDEFN
                     ETYTRRVFANDYREAEDHPQMVILQLLRFIFRLVWRRH"
BASE COUNT      113 a    116 c    109 g     85 t
ORIGIN


alignment_scores:
            Quality:  742.00             Length:    140
              Ratio:  5.300                Gaps:      0
 Percent Similarity: 100.000   Percent Identity: 100.000

alignment_block:
 US-09-508-832-4 x AF032460   ..

 Align seg 1/1  to: AF032460  from: 1  to: 423
```

```
      1 MetAlaLysGlnProSerAspValSerSerGluCysAspArgGluGlyGl 17
        ||||||||||||||||||||||||||||||||||||||||||||||||||
      1 ATGGCCAAGCAACCTTCTGATGTAAGTTCTGAGTGTGACAGAGAAGGTGG 50

     17 yGlnLeuGlnProAlaGluArgProProGlnLeuArgProGlyAlaProT 34
        ||||||||||||||||||||||||||||||||||||||||||||||||||
     51 ACAATTGCAGCCTGCTGAGAGGCCTCCCCAGCTCAGGCCTGGGGCCCCTA 100

     34 hrSerLeuGlnThrGluProGlnAspArgSerProAlaProMetSerCys 50
        ||||||||||||||||||||||||||||||||||||||||||||||||||
    101 CCTCCCTACAGACAGAACCGCAAGACAGGAGCCCGGCACCCATGAGTTGT 150

     51 AspLysSerThrGlnThrProSerProProCysGlnAlaPheAsnHisTy 67
        ||||||||||||||||||||||||||||||||||||||||||||||||||
    151 GACAAGTCAACACAAACCCCAAGTCCTCCTTGCCAGGCCTTCAACCACTA 200

     67 rLeuSerAlaMetAlaSerIleArgGlnSerGlnGluGluProGluAspL 84
        ||||||||||||||||||||||||||||||||||||||||||||||||||
    201 TCTCAGTGCAATGGCTTCCATACGACAGTCTCAGGAGGAACCTGAAGATC 250

     84 euArgProGluIleArgIleAlaGlnGluLeuArgArgIleGlyAspGlu 100
        ||||||||||||||||||||||||||||||||||||||||||||||||||
    251 TGCGCCCGGAGATACGGATTGCACAGGAGCTGCGGCGGATCGGAGACGAG 300

    101 PheAsnGluThrTyrThrArgArgValPheAlaAsnAspTyrArgGluAl 117
        ||||||||||||||||||||||||||||||||||||||||||||||||||
    301 TTCAACGAAACTTACACAAGGAGGGTGTTTGCAAATGATTACCGCGAGGC 350

    117 aGluAspHisProGlnMetValIleLeuGlnLeuLeuArgPheIlePheA 134
        ||||||||||||||||||||||||||||||||||||||||||||||||||
    351 TGAAGACCACCCTCAAATGGTTATCTTACAACTGTTACGCTTTATCTTCC 400

    134 rgLeuValTrpArgArgHis 140
        ||||||||||||||||||
    401 GTCTGGTATGGAGAAGGCAT 420
```

seq_name: gb_ro:AF136927

seq_documentation_block:

```
LOCUS       AF136927      423 bp    mRNA            ROD       21-APR-1999
DEFINITION  Rattus norvegicus Bcl-2 related apoptotic gene product BimL (bimL)
            mRNA, complete cds.
ACCESSION   AF136927
VERSION     AF136927.1  GI:4590514
KEYWORDS    .
SOURCE      Norway rat.
  ORGANISM  Rattus norvegicus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
            Rattus.
REFERENCE   1  (bases 1 to 423)
  AUTHORS   Chen,D., Simon,R.P. and Chen,J.
  TITLE     Cloning of rat bimEL and bimL, and their differential expression in
            ischemia and normal rat brain
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 423)
  AUTHORS   Chen,D., Simon,R.P. and Chen,J.
  TITLE     Direct Submission
  JOURNAL   Submitted (24-MAR-1999) Department of Neurology, BST, S-526,
            Pittsburgh University Medical School, 3500 Terrace Street,
            Pittsburgh, PA 15213, USA
FEATURES             Location/Qualifiers
     source          1. .423
                     /organism="Rattus norvegicus"
                     /db_xref="taxon:10116"
     gene            1. .423
                     /gene="bimL"
     CDS             1. .423
                     /gene="bimL"
                     /codon_start=1
                     /product="Bcl-2 related apoptotic gene product BimL"
                     /protein_id="AAD26594.1"
                     /db_xref="GI:4590515"
                     /translation="MAKQPSDVNSECDREGGQLQPAERPPQLRPGAPTSLQTESQDRS
                     PAPMSCDKSTQTPSPPCQAFNHYLSAMASIRQSQEEPEDLRPEIRIAQELRRIGDEFN
                     ETYTRRAFANDYREAEDHPQMVILQLLRFIFRLVWRRH"
BASE COUNT     116 a    113 c    112 g     82 t
ORIGIN
```

alignment_scores:

```
           Quality:  727.00            Length:    140
             Ratio:  5.230               Gaps:      0
 Percent Similarity:  99.286  Percent Identity: 97.857
```

alignment_block:

US-09-508-832-4 x AF136927 ..

Align seg 1/1 to: AF136927 from: 1 to: 423

```
      1 MetAlaLysGlnProSerAspValSerSerGluCysAspArgGluGlyGl 17
        ||||||||||||||||||||||||:::||||||||||||||||||||||||
      1 ATGGCCAAGCAACCTTCTGATGTAAATTCTGAGTGTGACAGAGAAGGTGG 50

     17 yGlnLeuGlnProAlaGluArgProProGlnLeuArgProGlyAlaProT 34
        ||||||||||||||||||||||||||||||||||||||||||||||||||
     51 ACAATTGCAGCCTGCTGAGAGGCCTCCCCAGCTCAGGCCTGGGGCCCCTA 100

     34 hrSerLeuGlnThrGluProGlnAspArgSerProAlaProMetSerCys 50
        ||||||||||||||||||  ||||||||||||||||||||||||||||||
    101 CCTCCCTACAGACAGAATCGCAAGACAGGAGCCCGGCACCCATGAGTTGT 150

     51 AspLysSerThrGlnThrProSerProProCysGlnAlaPheAsnHisTy 67
        ||||||||||||||||||||||||||||||||||||||||||||||||||
    151 GACAAGTCAACACAAACCCCAAGTCCTCCTTGCCAGGCCTTCAACCATTA 200

     67 rLeuSerAlaMetAlaSerIleArgGlnSerGlnGluGluProGluAspL 84
        ||||||||||||||||||||||||||||||||||||||||||||||||||
    201 TCTTAGCGCAATGGCTTCCATAAGGCAGTCTCAGGAGGAACCTGAAGATC 250

     84 euArgProGluIleArgIleAlaGlnGluLeuArgArgIleGlyAspGlu 100
        ||||||||||||||||||||||||||||||||||||||||||||||||||
    251 TGCGCCCAGAGATACGGATCGCACAGGAGCTGCGGAGGATCGGAGACGAG 300

    101 PheAsnGluThrTyrThrArgArgValPheAlaAsnAspTyrArgGluAl 117
        ||||||||||||||||||||||||:::||||||||||||||||||||||||
    301 TTCAATGAGACTTACACGAGGAGGGCGTTTGCAAACGATTACCGAGAGGC 350

    117 aGluAspHisProGlnMetValIleLeuGlnLeuLeuArgPheIlePheA 134
        ||||||||||||||||||||||||||||||||||||||||||||||||||
    351 GGAAGACCACCCGCAAATGGTTATCTTACAACTGTTACGATTCATCTTCC 400

    134 rgLeuValTrpArgArgHis 140
        ||||||||||||||||||
    401 GTCTGGTATGGAGAAGGCAT 420
```

seq_name: gb_pat:AX031283

seq_documentation_block:

```
LOCUS       AX031283      590 bp    DNA             PAT       20-SEP-2000
DEFINITION  Sequence 5 from Patent WO9914321.
ACCESSION   AX031283
VERSION     AX031283.1  GI:10278614
KEYWORDS    .
SOURCE      unidentified.
  ORGANISM  unidentified
            unclassified.
REFERENCE   1  (bases 1 to 590)
  AUTHORS   O'Reilly,L., Puthalakath,H., Adams,J., O'Connor,L., Cory,S.,
            Huang,D.C. and Strasser,A.
  TITLE     Novel therapeutic molecules
  JOURNAL   Patent: WO 9914321-A 5 25-MAR-1999;
            INST MEDICAL W & E HALL (AU) ; PUTHALAKATH HAMSA (AU) ; REILLY
            LORRAINE O (AU) ; ADAMS JERRY (AU) ; CONNOR LIAM O (AU) ; CORY
            SUZANNE (AU) ; HUANG DAVID C S (AU) ; STRASSER ANDREAS (AU)
```

```
FEATURES             Location/Qualifiers
     source          1. .590
                     /organism="unidentified"
                     /db_xref="taxon:32644"
     CDS             1. .>587
                     /note="unnamed protein product"
                     /codon_start=1
                     /protein_id="CAC09655.1"
                     /db_xref="GI:10278615"
                     /translation="MAKQPSDVSSECDREGGQLQPAERPPQLRPGAPTSLQTEPQGNP
                     DGEGDRCPHGSPQGPLAPPASPGPFATRSPLFIFVRRSSLLSRSSSGYFSFDTDRSPA
                     PMSCDKSTQTPSPPCQAFNHYLSAMASIRQSQEEPEDLRPEIRIAQELRRIGDEFNET
                     YTRRVFANDYREAEDHPQMVILQLLRFIFRLVWRR"
BASE COUNT      137 a    178 c    150 g    125 t
ORIGIN


alignment_scores:
           Quality:  704.00            Length:    196
             Ratio:  5.029               Gaps:      1
 Percent Similarity:  71.429   Percent Identity: 71.429

alignment_block:
 US-09-508-832-4 x AX031283   ..

 Align seg 1/1  to: AX031283  from: 1  to: 590

       1 MetAlaLysGlnProSerAspValSerSerGluCysAspArgGluGlyGl 17
         ||||||||||||||||||||||||||||||||||||||||||||||||||
       1 ATGGCCAAGCAACCTTCTGATGTAAGTTCTGAGTGTGACAGAGAAGGTGG 50

      17 yGlnLeuGlnProAlaGluArgProProGlnLeuArgProGlyAlaProT 34
         ||||||||||||||||||||||||||||||||||||||||||||||||||
      51 ACAATTGCAGCCTGCTGAGAGGCCTCCCCAGCTCAGGCCTGGGGCCCCTA 100

      34 hrSerLeuGlnThrGluProGln..................................... 41
         ||||||||||||||||||||||
     101 CCTCCCTACAGACAGAACCGCAAGGTAATCCCGACGGCGAAGGGGACCGC 150

      41 ..................................................... 41

     151 TGCCCCCACGGCAGCCCTCAGGGCCCGCTGGCCCCACCGGCCAGCCCTGG 200

      41 ..................................................... 41

     201 CCCTTTTGCTACCAGATCCCCACTTTTCATCTTTGTGAGAAGATCTTCTC 250

      42 ..........................................AspArgSer 44
                                                  |||||||||
     251 TGCTGTCCCGGTCCTCCAGTGGGTATTTCTCTTTTGACACAGACAGGAGC 300

      45 ProAlaProMetSerCysAspLysSerThrGlnThrProSerProProCy 61
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     301 CCGGCACCCATGAGTTGTGACAAGTCAACACAAACCCCAAGTCCTCCTTG 350

      61 sGlnAlaPheAsnHisTyrLeuSerAlaMetAlaSerIleArgGlnSerG 78
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     351 CCAGGCCTTCAACCACTATCTCAGTGCAATGGCTTCCATACGACAGTCTC 400

      78 lnGluGluProGluAspLeuArgProGluIleArgIleAlaGlnGluLeu 94
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     401 AGGAGGAACCTGAAGATCTGCGCCCGGAGATACGGATTGCACAGGAGCTG 450

      95 ArgArgIleGlyAspGluPheAsnGluThrTyrThrArgArgValPheAl 111
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     451 CGGCGGATCGGAGACGAGTTCAACGAAACTTACACAAGGAGGGTGTTTGC 500

     111 aAsnAspTyrArgGluAlaGluAspHisProGlnMetValIleLeuGlnL 128
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     501 AAATGATTACCGCGAGGCTGAAGACCACCCTCAAATGGTTATCTTACAAC 550

     128 euLeuArgPheIlePheArgLeuValTrpArgArgHis 140
         |||||||||||||||||||||||||||||||||||
     551 TGTTACGCTTTATCTTCCGTCTGGTATGGAGAAGGCAT 588

seq_name: gb_un:AX031309

seq_documentation_block:
LOCUS       AX031309      590 bp    DNA             UNA       20-SEP-2000
DEFINITION  Sequence 5 from Patent WO9914321.
ACCESSION   AX031309
VERSION     AX031309.1  GI:10278637
KEYWORDS    .
SOURCE      unidentified.
  ORGANISM  unidentified
            unclassified.
REFERENCE   1  (bases 1 to 590)
  AUTHORS   O'Reilly,L., Puthalakath,H., Adams,J., O'Connor,L., Cory,S.,
            Huang,D.C. and Strasser,A.
  TITLE     Novel therapeutic molecules
  JOURNAL   Patent: WO 9914321-A 25-MAR-1999;
            INST MEDICAL W & E HALL (AU) ; PUTHALAKATH HAMSA (AU) ; REILLY
            LORRAINE O (AU) ; ADAMS JERRY (AU) ; CONNOR LIAM O (AU) ; CORY
            SUZANNE (AU) ; HUANG DAVID C S (AU) ; STRASSER ANDREAS (AU)
FEATURES             Location/Qualifiers
     source          1. .590
                     /organism="unidentified"
                     /db_xref="taxon:32644"
     CDS             1. .>587
                     /note="unnamed protein product"
                     /codon_start=1
                     /protein_id="CAC09660.1"
                     /db_xref="GI:10278638"
                     /translation="MAKQPSDVSSECDREGGQLQPAERPPQLRPGAPTSLQTEPQGNP
                     DGEGDRCPHGSPQGPLAPPASPGPFATRSPLFIFVRRSSLLSRSSSGYFSFDTDRSPA
                     PMSCDKSTQTPSPPCQAFNHYLSAMASIRQSQEEPEDLRPEIRIAQELRRIGDEFNET
                     YTRRVFANDYREAEDHPQMVILQLLRFIFRLVWRR"
BASE COUNT      137 a    178 c    150 g    125 t
ORIGIN


alignment_scores:
           Quality:  704.00            Length:    196
             Ratio:  5.029               Gaps:      1
 Percent Similarity:  71.429   Percent Identity: 71.429

alignment_block:
 US-09-508-832-4 x AX031309   ..

 Align seg 1/1  to: AX031309  from: 1  to: 590

       1 MetAlaLysGlnProSerAspValSerSerGluCysAspArgGluGlyGl 17
         ||||||||||||||||||||||||||||||||||||||||||||||||||
       1 ATGGCCAAGCAACCTTCTGATGTAAGTTCTGAGTGTGACAGAGAAGGTGG 50

      17 yGlnLeuGlnProAlaGluArgProProGlnLeuArgProGlyAlaProT 34
         ||||||||||||||||||||||||||||||||||||||||||||||||||
      51 ACAATTGCAGCCTGCTGAGAGGCCTCCCCAGCTCAGGCCTGGGGCCCCTA 100

      34 hrSerLeuGlnThrGluProGln..................................... 41
         ||||||||||||||||||||||
     101 CCTCCCTACAGACAGAACCGCAAGGTAATCCCGACGGCGAAGGGGACCGC 150

      41 ..................................................... 41

     151 TGCCCCCACGGCAGCCCTCAGGGCCCGCTGGCCCCACCGGCCAGCCCTGG 200

      41 ..................................................... 41

     201 CCCTTTTGCTACCAGATCCCCACTTTTCATCTTTGTGAGAAGATCTTCTC 250

      42 ..........................................AspArgSer 44
                                                  |||||||||
     251 TGCTGTCCCGGTCCTCCAGTGGGTATTTCTCTTTTGACACAGACAGGAGC 300

      45 ProAlaProMetSerCysAspLysSerThrGlnThrProSerProProCy 61
```

```
          ||||||||||||||||||||||||||||||||||||||||||||||||||
     301 CCGGCACCCATGAGTTGTGACAAGTCAACACAAACCCCAAGTCCTCCTTG 350

      61 sGlnAlaPheAsnHisTyrLeuSerAlaMetAlaSerIleArgGlnSerG 78
          ||||||||||||||||||||||||||||||||||||||||||||||||||
     351 CCAGGCCTTCAACCACTATCTCAGTGCAATGGCTTCCATACGACAGTCTC 400

      78 lnGluGluProGluAspLeuArgProGluIleArgIleAlaGlnGluLeu 94
          ||||||||||||||||||||||||||||||||||||||||||||||||||
     401 AGGAGGAACCTGAAGATCTGCGCCCGGAGATACGGATTGCACAGGAGCTG 450

      95 ArgArgIleGlyAspGluPheAsnGluThrTyrThrArgArgValPheAl 111
          ||||||||||||||||||||||||||||||||||||||||||||||||||
     451 CGGCGGATCGGAGACGAGTTCAACGAAACTTACACAAGGAGGGTGTTTGC 500

     111 aAsnAspTyrArgGluAlaGluAspHisProGlnMetValIleLeuGlnL 128
          ||||||||||||||||||||||||||||||||||||||||||||||||||
     501 AAATGATTACCGCGAGGCTGAAGACCACCCTCAAATGGTTATCTTACAAC 550

     128 euLeuArgPheIlePheArgLeuValTrpArgArgHis 140
          ||||||||||||||||||||||||||||||||||||
     551 TGTTACGCTTTATCTTCCGTCTGGTATGGAGAAGGCAT 588

seq_name: gb_ro:AF032459

seq_documentation_block:
LOCUS       AF032459      591 bp    mRNA            ROD       19-FEB-1998
DEFINITION  Mus musculus BimEL mRNA, complete cds.
ACCESSION   AF032459
VERSION     AF032459.1  GI:2895499
KEYWORDS    .
SOURCE      house mouse.
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 591)
  AUTHORS   O'Connor,L., Strasser,A., O'Reilly,L.A., Hausmann,G., Adams,J.M.,
            Cory,S. and Huang,D.C.
  TITLE     Bim: a novel member of the Bcl-2 family that promotes apoptosis
  JOURNAL   EMBO J. 17 (2), 384-395 (1998)
  MEDLINE   98094360
   PUBMED   9430630
REFERENCE   2  (bases 1 to 591)
  AUTHORS   O'Connor,L., Strasser,A., O'Reilly,L.A., Hausmann,G., Adams,J.M.,
            Cory,S. and Huang,D.C.S.
  TITLE     Direct Submission
  JOURNAL   Submitted (03-NOV-1997) Molecular Genetics of Cancer, The Walter &
            Eliza Hall Institute of Medical Research, PO Royal Melbourne
            Hospital, Parkville, Victoria 3050, Australia
FEATURES             Location/Qualifiers
     source          1. .591
                     /organism="Mus musculus"
                     /db_xref="taxon:10090"
     CDS             1. .591
                     /note="pro-apoptotic BH3-containing Bcl-2 family member"
                     /codon_start=1
                     /product="BimEL"
                     /protein_id="AAC40029.1"
                     /db_xref="GI:2895500"
                     /translation="MAKQPSDVSSECDREGGQLQPAERPPQLRPGAPTSLQTEPQGNP
                     DGEGDRCPHGSPQGPLAPPASPGPFATRSPLFIFVRRSSLLSRSSSGYFSFDTDRSPA
                     PMSCDKSTQTPSPPCQAFNHYLSAMASIRQSQEEPEDLRPEIRIAQELRRIGDEFNET
                     YTRRVFANDYREAEDHPQMVILQLLRFIFRLVWRRH"
BASE COUNT      138 a    178 c    150 g    125 t
ORIGIN


alignment_scores:
            Quality:  704.00            Length:    196
              Ratio:  5.029               Gaps:      1
 Percent Similarity:  71.429   Percent Identity: 71.429

alignment_block:
```

```
US-09-508-832-4 x AF032459   ..

Align seg 1/1  to: AF032459  from: 1  to: 591

       1 MetAlaLysGlnProSerAspValSerSerGluCysAspArgGluGlyGl 17
         ||||||||||||||||||||||||||||||||||||||||||||||||||
       1 ATGGCCAAGCAACCTTCTGATGTAAGTTCTGAGTGTGACAGAGAAGGTGG 50

      17 yGlnLeuGlnProAlaGluArgProProGlnLeuArgProGlyAlaProT 34
         ||||||||||||||||||||||||||||||||||||||||||||||||||
      51 ACAATTGCAGCCTGCTGAGAGGCCTCCCCAGCTCAGGCCTGGGGCCCCTA 100

      34 hrSerLeuGlnThrGluProGln..................................... 41
         |||||||||||||||||||||||
     101 CCTCCCTACAGACAGAACCGCAAGGTAATCCCGACGGCGAAGGGGACCGC 150

      41 ..................................................... 41

     151 TGCCCCCACGGCAGCCCTCAGGGCCCGCTGGCCCCACCGGCCAGCCCTGG 200

      41 ..................................................... 41

     201 CCCTTTTGCTACCAGATCCCCACTTTTCATCTTTGTGAGAAGATCTTCTC 250

      42 ..........................................AspArgSer 44
                                                  ||||||||
     251 TGCTGTCCCGGTCCTCCAGTGGGTATTTCTCTTTTGACACAGACAGGAGC 300

      45 ProAlaProMetSerCysAspLysSerThrGlnThrProSerProProCy 61
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     301 CCGGCACCCATGAGTTGTGACAAGTCAACACAAACCCCAAGTCCTCCTTG 350

      61 sGlnAlaPheAsnHisTyrLeuSerAlaMetAlaSerIleArgGlnSerG 78
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     351 CCAGGCCTTCAACCACTATCTCAGTGCAATGGCTTCCATACGACAGTCTC 400

      78 lnGluGluProGluAspLeuArgProGluIleArgIleAlaGlnGluLeu 94
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     401 AGGAGGAACCTGAAGATCTGCGCCCGGAGATACGGATTGCACAGGAGCTG 450

      95 ArgArgIleGlyAspGluPheAsnGluThrTyrThrArgArgValPheAl 111
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     451 CGGCGGATCGGAGACGAGTTCAACGAAACTTACACAAGGAGGGTGTTTGC 500

     111 aAsnAspTyrArgGluAlaGluAspHisProGlnMetValIleLeuGlnL 128
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     501 AAATGATTACCGCGAGGCTGAAGACCACCCTCAAATGGTTATCTTACAAC 550

     128 euLeuArgPheIlePheArgLeuValTrpArgArgHis 140
         ||||||||||||||||||||||||||||||||||||
     551 TGTTACGCTTTATCTTCCGTCTGGTATGGAGAAGGCAT 588

seq_name: gb_ro:AF065433

seq_documentation_block:
LOCUS       AF065433      591 bp    mRNA            ROD       11-MAR-1999
DEFINITION  Rattus norvegicus Bcl-2 related ovarian death gene product BOD-L
            mRNA, complete cds.
ACCESSION   AF065433
VERSION     AF065433.1  GI:3228569
KEYWORDS    .
SOURCE      Norway rat.
  ORGANISM  Rattus norvegicus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
            Rattus.
REFERENCE   1  (bases 1 to 591)
  AUTHORS   Hsu,S.Y., Lin,P. and Hsueh,A.J.
  TITLE     BOD (Bcl-2-related ovarian death gene) is an ovarian BH3
            domain-containing proapoptotic Bcl-2 protein capable of
            dimerization with diverse antiapoptotic Bcl-2 members
  JOURNAL   Mol. Endocrinol. 12 (9), 1432-1440 (1998)
  MEDLINE   98400436
```

```
REFERENCE   2  (bases 1 to 591)
  AUTHORS   Hsu,S.Y. and Hsueh,A.J.W.
  TITLE     Direct Submission
  JOURNAL   Submitted (15-MAY-1998) GYN/OB, Stanford University, MSOB S385,
            Stanford, CA 94305, USA
FEATURES             Location/Qualifiers
     source          1. .591
                     /organism="Rattus norvegicus"
                     /db_xref="taxon:10116"
     CDS             1. .591
                     /codon_start=1
                     /product="Bcl-2 related ovarian death gene product BOD-L"
                     /protein_id="AAC23595.1"
                     /db_xref="GI:3228570"
                     /translation="MAKQPSDVNSECDREGGQLQPAERPPQLRPGAPTSLQTESQGNP
                     DGEGDRCPHGSPQGPLAPPASPGPFATRSPLFIFVRRSSLLSRSSSGYFSFDTDRSPA
                     PMSCDKSTQTPSPPCQAFNHYLSAMASIRQSQEEPEDLRPEIRIAQELRRIGDEFNET
                     YTRRAFANDYREAEDHPQMVILQLLRFIFRLVWRRH"
BASE COUNT      139 a    177 c    153 g    122 t
ORIGIN


alignment_scores:
            Quality:  689.00            Length:    196
              Ratio:  4.957               Gaps:      1
 Percent Similarity:  70.918   Percent Identity: 69.898

alignment_block:
 US-09-508-832-4 x AF065433   ..

 Align seg 1/1  to: AF065433  from: 1  to: 591

      -1 MetAlaLysGlnProSerAspValSerSerGluCysAspArgGluGlyGl 17
         ||||||||||||||||||||||:::||||||||||||||||||||||||||
          1 ATGGCCAAGCAACCTTCTGATGTAAATTCTGAGTGTGACAGAGAAGGTGG 50

      17 yGlnLeuGlnProAlaGluArgProProGlnLeuArgProGlyAlaProT 34
         ||||||||||||||||||||||||||||||||||||||||||||||||||||
      51 ACAATTGCAGCCTGCTGAGAGGCCTCCCCAGCTCAGGCCTGGGGCCCCTA 100

      34 hrSerLeuGlnThrGluProGln..................... 41
         ||||||||||||||||   |||
     101 CCTCCCTACAGACAGAATCGCAAGGTAATCCCGACGGCGAAGGGGACCGC 150

      41 ............................................. 41

     151 TGCCCCCACGGCAGCCCTCAGGGCCCGCTGGCCCCACCGGCCAGCCCTGG 200

      41 ............................................. 41

     201 TCCTTTTGCTACCAGATCCCCACTTTTCATCTTTGTGAGAAGATCTTCTC 250

      42 .....................................AspArgSer 44
                                              ||||||||
     251 TGCTGTCCCGGTCCTCCAGTGGGTATTTCTCTTTTGACACAGACAGGAGC 300

      45 ProAlaProMetSerCysAspLysSerThrGlnThrProSerProProCy 61
         ||||||||||||||||||||||||||||||||||||||||||||||||||||
     301 CCGGCACCCATGAGTTGTGACAAGTCAACACAAACCCCAAGTCCTCCTTG 350

      61 sGlnAlaPheAsnHisTyrLeuSerAlaMetAlaSerIleArgGlnSerG 78
         ||||||||||||||||||||||||||||||||||||||||||||||||||||
     351 CCAGGCCTTCAACCATTATCTCAGTGCAATGGCTTCCATAAGGCAGTCTC 400

      78 lnGluGluProGluAspLeuArgProGluIleArgIleAlaGlnGluLeu 94
         ||||||||||||||||||||||||||||||||||||||||||||||||||||
     401 AGGAGGAACCTGAAGATCTGCGCCCAGAGATACGGATCGCACAGGAGCTG 450

      95 ArgArgIleGlyAspGluPheAsnGluThrTyrThrArgArgValPheAl 111
         ||||||||||||||||||||||||||||||||||||||||||:::||||||
     451 CGGCGGATCGGAGACGAGTTCAATGAGACTTACACGAGGAGGGCGTTTGC 500

     111 aAsnAspTyrArgGluAlaGluAspHisProGlnMetValIleLeuGlnL 128
         ||||||||||||||||||||||||||||||||||||||||||||||||||||
     501 AAACGATTACCGAGAGGCGGAAGACCACCCGCAAATGGTTATCTTACAAC 550

     128 euLeuArgPheIlePheArgLeuValTrpArgArgHis 140
         ||||||||||||||||||||||||||||||||||||
     551 TGTTACGATTCATCTTCCGTCTGGTCTGGAGAAGGCAC 588

seq_name: gb_pat:AX031285

seq_documentation_block:
LOCUS       AX031285      416 bp    DNA             PAT       20-SEP-2000
DEFINITION  Sequence 7 from Patent WO9914321.
ACCESSION   AX031285
VERSION     AX031285.1  GI:10278616
KEYWORDS    .
SOURCE      unidentified.
  ORGANISM  unidentified
            unclassified.
REFERENCE   1  (bases 1 to 416)
  AUTHORS   O'Reilly,L., Puthalakath,H., Adams,J., O'Connor,L., Cory,S.,
            Huang,D.C. and Strasser,A.
  TITLE     Novel therapeutic molecules
  JOURNAL   Patent: WO 9914321-A 7 25-MAR-1999;
            INST MEDICAL W & E HALL (AU) ; PUTHALAKATH HAMSA (AU) ; REILLY
            LORRAINE O (AU) ; ADAMS JERRY (AU) ; CONNOR LIAM O (AU) ; CORY
            SUZANNE (AU) ; HUANG DAVID C S (AU) ; STRASSER ANDREAS (AU)
FEATURES             Location/Qualifiers
     source          1. .416
                     /organism="unidentified"
                     /db_xref="taxon:32644"
     CDS             1. .>414
                     /note="unnamed protein product"
                     /codon_start=1
                     /protein_id="CAC09656.1"
                     /db_xref="GI:10278617"
                     /translation="MAKQPSDVSSECDREGRQLQPAERPPQLRPGAPTSLQTEPQDRS
                     PAPMSCDKSTQTPSPPCQAFNHYLSAMASMRQAEPADMRPEIWIAQELRRIGDEFNAY
                     YARRVFLNNYQAAEDHPRMVILRLLRYIVRLVWRMH"
BASE COUNT      113 a    113 c    103 g     87 t
ORIGIN


alignment_scores:
            Quality:  629.00            Length:    140
              Ratio:  4.876               Gaps:      1
 Percent Similarity:  92.143   Percent Identity: 85.714

alignment_block:
 US-09-508-832-4 x AX031285   ..

 Align seg 1/1  to: AX031285  from: 1  to: 416

       1 MetAlaLysGlnProSerAspValSerSerGluCysAspArgGluGlyGl 17
         ||||||||||||||||||||||||||||||||||||||||||||||||||||
       1 ATGGCAAAGCAACCTTCTGATGTAAGTTCTGAGTGTGACCGAGAAGGTAG 50

      17 yGlnLeuGlnProAlaGluArgProProGlnLeuArgProGlyAlaProT 34
         ||||||||||||||||||||||||||||||||||||||||||||||||||||
      51 ACAATTGCAGCCTGCGGAGAGGCCTCCCCAGCTCAGACCTGGGGCCCCTA 100

      34 hrSerLeuGlnThrGluProGlnAspArgSerProAlaProMetSerCys 50
         ||||||||||||||||||||||||||||||||||||||||||||||||||||
     101 CCTCCCTACAGACAGAGCCACAAGACAGGAGCCCAGCACCCATGAGTTGT 150

      51 AspLysSerThrGlnThrProSerProProCysGlnAlaPheAsnHisTy 67
         ||||||||||||||||||||||||||||||||||||||||||||||||||||
     151 GACAAATCAACACAAACCCCAAGTCCTCCTTGCCAGGCCTTCAACCACTA 200

      67 rLeuSerAlaMetAlaSerIleArgGlnSerGlnGluGluProGluAspL 84
         ||||||||||||||||||:::||||||:::      ||||||   |||:
     201 TCTCAGTGCAATGGCTTCCATGAGGCAGGCT......GAACCTGCAGATA 244

      84 euArgProGluIleArgIleAlaGlnGluLeuArgArgIleGlyAspGlu 100
```

```
     ::||||||||||||||   ||||||||||||||||||||||||||||||||
    245 TGCGCCCAGAGATATGGATCGCCCAAGAGTTGCGGCGTATCGGAGACGAG 294

    101 PheAsnGluThrTyrThrArgArgValPheAlaAsnAspTyrArgGluAl 117
        ||||||      |||:::||||||||||||||   |||:::|||:::   ||
    295 TTTAACGCTTACTATGCAAGGAGGGTATTTTTGAATAATTACCAAGCAGC 344

    117 aGluAspHisProGlnMetValIleLeuGlnLeuLeuArgPheIlePheA 134
        ||||||||||||||:::||||||||||||:::||||||||||:::|||   |
    345 CGAAGACCACCCACGAATGGTTATCTTACGACTGTTACGTTACATTGTCC 394

    134 rgLeuValTrpArgArgHis 140
        ||||||||||||||   |||
    395 GCCTGGTGTGGAGAATGCAT 414

seq_name: gb_un:AX031311

seq_documentation_block:
LOCUS       AX031311      416 bp    DNA             UNA       20-SEP-2000
DEFINITION  Sequence 7 from Patent WO9914321.
ACCESSION   AX031311
VERSION     AX031311.1  GI:10278639
KEYWORDS    .
SOURCE      unidentified.
  ORGANISM  unidentified
            unclassified.
REFERENCE   1  (bases 1 to 416)
  AUTHORS   O'Reilly,L., Puthalakath,H., Adams,J., O'Connor,L., Cory,S.,
            Huang,D.C. and Strasser,A.
  TITLE     Novel therapeutic molecules
  JOURNAL   Patent: WO 9914321-A 25-MAR-1999;
            INST MEDICAL W & E HALL (AU) ; PUTHALAKATH HAMSA (AU) ; REILLY
            LORRAINE O (AU) ; ADAMS JERRY (AU) ; CONNOR LIAM O (AU) ; CORY
            SUZANNE (AU) ; HUANG DAVID C S (AU) ; STRASSER ANDREAS (AU)
FEATURES             Location/Qualifiers
     source          1. .416
                     /organism="unidentified"
                     /db_xref="taxon:32644"
     CDS             1. .>414
                     /note="unnamed protein product"
                     /codon_start=1
                     /protein_id="CAC09661.1"
                     /db_xref="GI:10278640"
                     /translation="MAKQPSDVSSECDREGRQLQPAERPPQLRPGAPTSLQTEPQDRS
                     PAPMSCDKSTQTPSPPCQAFNHYLSAMASMRQAEPADMRPEIWIAQELRRIGDEFNAY
                     YARRVFLNNYQAAEDHPRMVILRLLRYIVRLVWRMH"
BASE COUNT      113 a    113 c    103 g    87 t
ORIGIN


alignment_scores:
            Quality:  629.00            Length:    140
              Ratio:  4.876               Gaps:      1
 Percent Similarity:  92.143   Percent Identity: 85.714

alignment_block:
 US-09-508-832-4 x AX031311   ..

 Align seg 1/1  to: AX031311  from: 1  to: 416

       1 MetAlaLysGlnProSerAspValSerSerGluCysAspArgGluGlyGl 17
         ||||||||||||||||||||||||||||||||||||||||||||||||||
       1 ATGGCAAAGCAACCTTCTGATGTAAGTTCTGAGTGTGACCGAGAAGGTAG 50

      17 yGlnLeuGlnProAlaGluArgProProGlnLeuArgProGlyAlaProT 34
         |||||||||||||||||||||||||||||||||||||||||||||||||
      51 ACAATTGCAGCCTGCGGAGAGGCCTCCCCAGCTCAGACCTGGGGCCCCTA 100

      34 hrSerLeuGlnThrGluProGlnAspArgSerProAlaProMetSerCys 50
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     101 CCTCCCTACAGACAGAGCCACAAGACAGGAGCCCAGCACCCATGAGTTGT 150

      51 AspLysSerThrGlnThrProSerProProCysGlnAlaPheAsnHisTy 67

         ||||||||||||||||||||||||||||||||||||||||||||||||||
     151 GACAAATCAACACAAACCCCAAGTCCTCCTTGCCAGGCCTTCAACCACTA 200

      67 rLeuSerAlaMetAlaSerIleArgGlnSerGlnGluGluProGluAspL 84
         ||||||||||||||||||:::||||||:::      ||||||   |||:
     201 TCTCAGTGCAATGGCTTCCATGAGGCAGGCT......GAACCTGCAGATA 244

      84 euArgProGluIleArgIleAlaGlnGluLeuArgArgIleGlyAspGlu 100
         ::|||||||||||||   ||||||||||||||||||||||||||||||||
     245 TGCGCCCAGAGATATGGATCGCCCAAGAGTTGCGGCGTATCGGAGACGAG 294

     101 PheAsnGluThrTyrThrArgArgValPheAlaAsnAspTyrArgGluAl 117
         ||||||      |||:::||||||||||||||   |||:::|||:::   ||
     295 TTTAACGCTTACTATGCAAGGAGGGTATTTTTGAATAATTACCAAGCAGC 344

     117 aGluAspHisProGlnMetValIleLeuGlnLeuLeuArgPheIlePheA 134
         ||||||||||||||:::||||||||||||:::||||||||||:::|||   |
     345 CGAAGACCACCCACGAATGGTTATCTTACGACTGTTACGTTACATTGTCC 394

     134 rgLeuValTrpArgArgHis 140
         ||||||||||||||   |||
     395 GCCTGGTGTGGAGAATGCAT 414

seq_name: gb_pr:AF032458

seq_documentation_block:
LOCUS       AF032458      417 bp    mRNA            PRI       19-FEB-1998
DEFINITION  Homo sapiens BimL mRNA, complete cds.
ACCESSION   AF032458
VERSION     AF032458.1  GI:2895497
KEYWORDS    .
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 417)
  AUTHORS   O'Connor,L., Strasser,A., O'Reilly,L.A., Hausmann,G., Adams,J.M.,
            Cory,S. and Huang,D.C.
  TITLE     Bim: a novel member of the Bcl-2 family that promotes apoptosis
  JOURNAL   EMBO J. 17 (2), 384-395 (1998)
  MEDLINE   98094360
   PUBMED   9430630
REFERENCE   2  (bases 1 to 417)
  AUTHORS   O'Connor,L., Strasser,A., O'Reilly,L.A., Hausmann,G., Adams,J.M.,
            Cory,S. and Huang,D.C.S.
  TITLE     Direct Submission
  JOURNAL   Submitted (03-NOV-1997) Molecular Genetics of Cancer, The Walter &
            Eliza Hall Institute of Medical Research, PO Royal Melbourne
            Hospital, Parkville, Victoria 3050, Australia
FEATURES             Location/Qualifiers
     source          1. .417
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
     CDS             1. .417
                     /note="pro-apoptotic BH3-containing Bcl-2 family member"
                     /codon_start=1
                     /product="BimL"
                     /protein_id="AAC39594.1"
                     /db_xref="GI:2895498"
                     /translation="MAKQPSDVSSECDREGRQLQPAERPPQLRPGAPTSLQTEPQDRS
                     PAPMSCDKSTQTPSPPCQAFNHYLSAMASMRQAEPADMRPEIWIAQELRRIGDEFNAY
                     YARRVFLNNYQAAEDHPRMVILRLLRYIVRLVWRMH"
BASE COUNT      114 a    113 c    103 g    87 t
ORIGIN


alignment_scores:
            Quality:  629.00            Length:    140
              Ratio:  4.876               Gaps:      1
 Percent Similarity:  92.143   Percent Identity: 85.714

alignment_block:
 US-09-508-832-4 x AF032458   ..
```

```
 Align seg 1/1  to: AF032458  from: 1  to: 417

       1 MetAlaLysGlnProSerAspValSerSerGluCysAspArgGluGlyGl 17
         ||||||||||||||||||||||||||||||||||||||||||||||||||
       1 ATGGCAAAGCAACCTTCTGATGTAAGTTCTGAGTGTGACCGAGAAGGTAG 50

      17 yGlnLeuGlnProAlaGluArgProProGlnLeuArgProGlyAlaProT 34
         ||||||||||||||||||||||||||||||||||||||||||||||||||
      51 ACAATTGCAGCCTGCGGAGAGGCCTCCCCAGCTCAGACCTGGGGCCCCTA 100

      34 hrSerLeuGlnThrGluProGlnAspArgSerProAlaProMetSerCys 50
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     101 CCTCCCTACAGACAGAGCCACAAGACAGGAGCCCAGCACCCATGAGTTGT 150

      51 AspLysSerThrGlnThrProSerProProCysGlnAlaPheAsnHisTy 67
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     151 GACAAATCAACACAAACCCCAAGTCCTCCTTGCCAGGCCTTCAACCACTA 200

      67 rLeuSerAlaMetAlaSerIleArgGlnSerGlnGluGluProGluAspL 84
         ||||||||||||||||||:::||||||:::      ||||||   |||:
     201 TCTCAGTGCAATGGCTTCCATGAGGCAGGCT......GAACCTGCAGATA 244

      84 euArgProGluIleArgIleAlaGlnGluLeuArgArgIleGlyAspGlu 100
         ::|||||||||||   ||||||||||||||||||||||||||||||||||
     245 TGCGCCCAGAGATATGGATCGCCCAAGAGTTGCGGCGTATCGGAGACGAG 294

     101 PheAsnGluThrTyrThrArgArgValPheAlaAsnAspTyrArgGluAl 117
         ||||||      |||:::||||||||||||   |||:::|||:::   ||
     295 TTTAACGCTTACTATGCAAGGAGGGTATTTTTGAATAATTACCAAGCAGC 344

     117 aGluAspHisProGlnMetValIleLeuGlnLeuLeuArgPheIlePheA 134
         |||||||||||||:::|||||||||||||:::||||||||||:::|||   |
     345 CGAAGACCACCCACGAATGGTTATCTTACGACTGTTACGTTACATTGTCC 394

     134 rgLeuValTrpArgArgHis 140
         ||||||||||||||   |||
     395 GCCTGGTGTGGAGAATGCAT 414

seq_name: gb_pat:AX031287

seq_documentation_block:
LOCUS       AX031287      596 bp    DNA             PAT       20-SEP-2000
DEFINITION  Sequence 9 from Patent WO9914321.
ACCESSION   AX031287
VERSION     AX031287.1  GI:10278618
KEYWORDS    .
SOURCE      unidentified.
  ORGANISM  unidentified
            unclassified.
REFERENCE   1  (bases 1 to 596)
  AUTHORS   O'Reilly,L., Puthalakath,H., Adams,J., O'Connor,L., Cory,S.,
            Huang,D.C. and Strasser,A.
  TITLE     Novel therapeutic molecules
  JOURNAL   Patent: WO 9914321-A 9 25-MAR-1999;
            INST MEDICAL W & E HALL (AU) ; PUTHALAKATH HAMSA (AU) ; REILLY
            LORRAINE O (AU) ; ADAMS JERRY (AU) ; CONNOR LIAM O (AU) ; CORY
            SUZANNE (AU) ; HUANG DAVID C S (AU) ; STRASSER ANDREAS (AU)
FEATURES             Location/Qualifiers
     source          1. .596
                     /organism="unidentified"
                     /db_xref="taxon:32644"
     CDS             1. .>594
                     /note="unnamed protein product"
                     /codon_start=1
                     /protein_id="CAC09657.1"
                     /db_xref="GI:10278619"
                     /translation="MAKQPSDVSSECDREGRQLQPAERPPQLRPGAPTSLQTEPQGNP
                     EGNHGGEGDSCPHGSPQGPLAPPASPGPFATRSPLFIFMRRSSLLSRSSSGYFSFDTD
                     RSPAPMSCDKSTQTPSPPCQAFNHYLSAMASMRQAEPADMRPEIWIAQELRRIGDEFN
                     AYYARRVFLNNYQAAEDHPRMVILRLLRYIVRLVWRMH"
BASE COUNT      145 a    175 c    146 g    130 t
ORIGIN
```

```
alignment_scores:
          Quality:  589.00              Length:    200
            Ratio:  4.566                Gaps:      2
 Percent Similarity:  64.500   Percent Identity: 60.000

alignment_block:
US-09-508-832-4 x AX031287   ..

 Align seg 1/1  to: AX031287  from: 1  to: 596

       1 MetAlaLysGlnProSerAspValSerSerGluCysAspArgGluGlyGl 17
         ||||||||||||||||||||||||||||||||||||||||||||||||||
       1 ATGGCAAAGCAACCTTCTGATGTAAGTTCTGAGTGTGACCGAGAAGGTAG 50

      17 yGlnLeuGlnProAlaGluArgProProGlnLeuArgProGlyAlaProT 34
         ||||||||||||||||||||||||||||||||||||||||||||||||||
      51 ACAATTGCAGCCTGCGGAGAGGCCTCCCCAGCTCAGACCTGGGGCCCCTA 100

      34 hrSerLeuGlnThrGluProGln..........................  41
         ||||||||||||||||||||||||
     101 CCTCCCTACAGACAGAGCCACAAGGTAATCCTGAAGGCAATCACGGAGGT 150

      41 ..................................................  41

     151 GAAGGGGACAGCTGCCCCCACGGCAGCCCTCAGGGCCCGCTGGCCCCACC 200

      41 ..................................................  41

     201 TGCCAGCCCTGGCCCTTTTGCTACCAGATCCCCGCTTTTCATCTTTATGA 250

      41 ..................................................  41

     251 GAAGATCCTCCCTGCTGTCTCGATCCTCCAGTGGGTATTTCTCTTTTGAC 300

      42 ...AspArgSerProAlaProMetSerCysAspLysSerThrGlnThrPr 57
            ||||||||||||||||||||||||||||||||||||||||||||||||
     301 ACAGACAGGAGCCCAGCACCCATGAGTTGTGACAAATCAACACAAACCCC 350

      57 oSerProProCysGlnAlaPheAsnHisTyrLeuSerAlaMetAlaSerI 74
         ||||||||||||||||||||||||||||||||||||||||||||||||||:
     351 AAGTCCTCCTTGCCAGGCCTTCAACCACTATCTCAGTGCAATGGCTTCCA 400

      74 leArgGlnSerGlnGluGluProGluAspLeuArgProGluIleArgIle 90
         ::||||||:::      ||||||   |||:::||||||||||||   |||
     401 TGAGGCAGGCT......GAACCTGCAGATATGCGCCCAGAGATATGGATC 444

      91 AlaGlnGluLeuArgArgIleGlyAspGluPheAsnGluThrTyrThrAr 107
         ||||||||||||||||||||||||||||||||||||      |||:::||
     445 GCCCAAGAGTTGCGGCGTATCGGAGACGAGTTTAACGCTTACTATGCAAG 494

     107 gArgValPheAlaAsnAspTyrArgGluAlaGluAspHisProGlnMetV 124
         ||||||||||   |||:::|||:::   ||||||||||||||:::||||
     495 GAGGGTATTTTTGAATAATTACCAAGCAGCCGAAGACCACCCACGAATGG 544

     124 alIleLeuGlnLeuLeuArgPheIlePheArgLeuValTrpArgArgHis 140
         |||||||:::||||||||:::|||   ||||||||||||||   |||
     545 TTATCTTACGACTGTTACGTTACATTGTCCGCCTGGTGTGGAGAATGCAT 594

seq_name: gb_un:AX031313

seq_documentation_block:
LOCUS       AX031313      596 bp    DNA             UNA       20-SEP-2000
DEFINITION  Sequence 9 from Patent WO9914321.
ACCESSION   AX031313
VERSION     AX031313.1  GI:10278641
KEYWORDS    .
SOURCE      unidentified.
  ORGANISM  unidentified
            unclassified.
REFERENCE   1  (bases 1 to 596)
  AUTHORS   O'Reilly,L., Puthalakath,H., Adams,J., O'Connor,L., Cory,S.,
```

```
          Huang,D.C. and Strasser,A.
  TITLE   Novel therapeutic molecules
  JOURNAL Patent: WO 9914321-A 25-MAR-1999;
          INST MEDICAL W & E HALL (AU) ; PUTHALAKATH HAMSA (AU) ; REILLY
          LORRAINE O (AU) ; ADAMS JERRY (AU) ; CONNOR LIAM O (AU) ; CORY
          SUZANNE (AU) ; HUANG DAVID C S (AU) ; STRASSER ANDREAS (AU)
FEATURES             Location/Qualifiers
     source          1. .596
                     /organism="unidentified"
                     /db_xref="taxon:32644"
     CDS             1. .>594
                     /note="unnamed protein product"
                     /codon_start=1
                     /protein_id="CAC09662.1"
                     /db_xref="GI:10278642"
                     /translation="MAKQPSDVSSECDREGRQLQPAERPPQLRPGAPTSLQTEPQGNP
                     EGNHGGEGDSCPHGSPQGPLAPPASPGPFATRSPLFIFMRRSSLLSRSSSGYFSFDTD
                     RSPAPMSCDKSTQTPSPPCQAFNHYLSAMASMRQAEPADMRPEIWIAQELRRIGDEFN
                     AYYARRVFLNNYQAAEDHPRMVILRLLRYIVRLVWRMH"
BASE COUNT      145 a    175 c    146 g    130 t
ORIGIN


alignment_scores:
            Quality:  589.00            Length:    200
              Ratio:  4.566               Gaps:      2
 Percent Similarity:  64.500   Percent Identity: 60.000

alignment_block:
 US-09-508-832-4 x AX031313   ..

 Align seg 1/1  to: AX031313  from: 1  to: 596

       1 MetAlaLysGlnProSerAspValSerSerGluCysAspArgGluGlyGl 17
         ||||||||||||||||||||||||||||||||||||||||||||||||||
       1 ATGGCAAAGCAACCTTCTGATGTAAGTTCTGAGTGTGACCGAGAAGGTAG 50

      17 yGlnLeuGlnProAlaGluArgProProGlnLeuArgProGlyAlaProT 34
         ||||||||||||||||||||||||||||||||||||||||||||||||||
      51 ACAATTGCAGCCTGCGGAGAGGCCTCCCCAGCTCAGACCTGGGGCCCCTA 100

      34 hrSerLeuGlnThrGluProGln..........................  41
         |||||||||||||||||||||||
     101 CCTCCCTACAGACAGAGCCACAAGGTAATCCTGAAGGCAATCACGGAGGT 150

      41 .....................................................  41

     151 GAAGGGGACAGCTGCCCCCACGGCAGCCCTCAGGGCCCGCTGGCCCCACC 200

      41 .....................................................  41

     201 TGCCAGCCCTGGCCCTTTTGCTACCAGATCCCCGCTTTTCATCTTTATGA 250

      41 .....................................................  41

     251 GAAGATCCTCCCTGCTGTCTCGATCCTCCAGTGGGTATTTCTCTTTTGAC 300

      42 ...AspArgSerProAlaProMetSerCysAspLysSerThrGlnThrPr  57
            ||||||||||||||||||||||||||||||||||||||||||||||||
     301 ACAGACAGGAGCCCAGCACCCATGAGTTGTGACAAATCAACACAAACCCC 350

      57 oSerProProCysGlnAlaPheAsnHisTyrLeuSerAlaMetAlaSerI 74
         ||||||||||||||||||||||||||||||||||||||||||||||||:
     351 AAGTCCTCCTTGCCAGGCCTTCAACCACTATCTCAGTGCAATGGCTTCCA 400

      74 leArgGlnSerGlnGluGluProGluAspLeuArgProGluIleArgIle 90
         ::||||||:::      ||||||   |||:::||||||||||   |||
     401 TGAGGCAGGCT......GAACCTGCAGATATGCGCCCAGAGATATGGATC 444

      91 AlaGlnGluLeuArgArgIleGlyAspGluPheAsnGluThrTyrThrAr 107
         ||||||||||||||||||||||||||||||||      |||:::||
     445 GCCCAAGAGTTGCGGCGTATCGGAGACGAGTTTAACGCTTACTATGCAAG 494

     107 gArgValPheAlaAsnAspTyrArgGluAlaGluAspHisProGlnMetV 124
         |||||||||   |||:::|||:::   ||||||||||||||||:::||||
     495 GAGGGTATTTTTGAATAATTACCAAGCAGCCGAAGACCACCCACGAATGG 544

     124 alIleLeuGlnLeuLeuArgPheIlePheArgLeuValTrpArgArgHis 140
         ||||||||:::|||||||||:::|||   |||||||||||||||   |||
     545 TTATCTTACGACTGTTACGTTACATTGTCCGCCTGGTGTGGAGAATGCAT 594

seq_name: gb_pr:AF032457

seq_documentation_block:
LOCUS       AF032457      597 bp    mRNA            PRI       19-FEB-1998
DEFINITION  Homo sapiens BimEL mRNA, complete cds.
ACCESSION   AF032457
VERSION     AF032457.1  GI:2895495
KEYWORDS    .
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 597)
  AUTHORS   O'Connor,L., Strasser,A., O'Reilly,L.A., Hausmann,G., Adams,J.M.,
            Cory,S. and Huang,D.C.
  TITLE     Bim: a novel member of the Bcl-2 family that promotes apoptosis
  JOURNAL   EMBO J. 17 (2), 384-395 (1998)
  MEDLINE   98094360
   PUBMED   9430630
REFERENCE   2  (bases 1 to 597)
  AUTHORS   O'Connor,L., Strasser,A., O'Reilly,L.A., Hausmann,G., Adams,J.M.,
            Cory,S. and Huang,D.C.S.
  TITLE     Direct Submission
  JOURNAL   Submitted (03-NOV-1997) Molecular Genetics of Cancer, The Walter &
            Eliza Hall Institute of Medical Research, PO Royal Melbourne
            Hospital, Parkville, Victoria 3050, Australia
FEATURES             Location/Qualifiers
     source          1. .597
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
     CDS             1. .597
                     /note="pro-apoptotic BH3-containing Bcl-2 family member"
                     /codon_start=1
                     /product="BimEL"
                     /protein_id="AAC39593.1"
                     /db_xref="GI:2895496"
                     /translation="MAKQPSDVSSECDREGRQLQPAERPPQLRPGAPTSLQTEPQGNP
                     EGNHGGEGDSCPHGSPQGPLAPPASPGPFATRSPLFIFMRRSSLLSRSSSGYFSFDTD
                     RSPAPMSCDKSTQTPSPPCQAFNHYLSAMASMRQAEPADMRPEIWIAQELRRIGDEFN
                     AYYARRVFLNNYQAAEDHPRMVILRLLRYIVRLVWRMH"
BASE COUNT      146 a    175 c    146 g    130 t
ORIGIN


alignment_scores:
            Quality:  589.00            Length:    200
              Ratio:  4.566               Gaps:      2
 Percent Similarity:  64.500   Percent Identity: 60.000

alignment_block:
 US-09-508-832-4 x AF032457   ..

 Align seg 1/1  to: AF032457  from: 1  to: 597

       1 MetAlaLysGlnProSerAspValSerSerGluCysAspArgGluGlyGl 17
         ||||||||||||||||||||||||||||||||||||||||||||||||||
       1 ATGGCAAAGCAACCTTCTGATGTAAGTTCTGAGTGTGACCGAGAAGGTAG 50

      17 yGlnLeuGlnProAlaGluArgProProGlnLeuArgProGlyAlaProT 34
         ||||||||||||||||||||||||||||||||||||||||||||||||||
      51 ACAATTGCAGCCTGCGGAGAGGCCTCCCCAGCTCAGACCTGGGGCCCCTA 100

      34 hrSerLeuGlnThrGluProGln..........................  41
         |||||||||||||||||||||||
     101 CCTCCCTACAGACAGAGCCACAAGGTAATCCTGAAGGCAATCACGGAGGT 150
```

```
     41 ..................................................  41

    151 GAAGGGGACAGCTGCCCCCACGGCAGCCCTCAGGGCCCGCTGGCCCCACC 200

     41 ..................................................  41

    201 TGCCAGCCCTGGCCCTTTTGCTACCAGATCCCCGCTTTTCATCTTTATGA 250

     41 ..................................................  41

    251 GAAGATCCTCCCTGCTGTCTCGATCCTCCAGTGGGTATTTCTCTTTTGAC 300

     42 ...AspArgSerProAlaProMetSerCysAspLysSerThrGlnThrPr  57
         ||||||||||||||||||||||||||||||||||||||||||||||||||
    301 ACAGACAGGAGCCCAGCACCCATGAGTTGTGACAAATCAACACAAACCCC 350

     57 oSerProProCysGlnAlaPheAsnHisTyrLeuSerAlaMetAlaSerI  74
         ||||||||||||||||||||||||||||||||||||||||||||||||:
    351 AAGTCCTCCTTGCCAGGCCTTCAACCACTATCTCAGTGCAATGGCTTCCA 400

     74 leArgGlnSerGlnGluGluProGluAspLeuArgProGluIleArgIle  90
         ::|||||:::      ||||||   |||:::|||||||||||||   |||
    401 TGAGGCAGGCT......GAACCTGCAGATATGCGCCCAGAGATATGGATC 444

     91 AlaGlnGluLeuArgArgIleGlyAspGluPheAsnGluThrTyrThrAr 107
         ||||||||||||||||||||||||||||||||      |||:::||
    445 GCCCAAGAGTTGCGGCGTATCGGAGACGAGTTTAACGCTTACTATGCAAG 494

    107 gArgValPheAlaAsnAspTyrArgGluAlaGluAspHisProGlnMetV 124
         ||||||||||   |||:::|||:::   |||||||||||||||:::||||
    495 GAGGGTATTTTTGAATAATTACCAAGCAGCCGAAGACCACCCACGAATGG 544

    124 alIleLeuGlnLeuLeuArgPheIlePheArgLeuValTrpArgArgHis 140
         ||||||||:::|||||||||:::|||   ||||||||||||||   |||
    545 TTATCTTACGACTGTTACGTTACATTGTCCGCCTGGTGTGGAGAATGCAT 594

seq_name: gb_pat:AX031279

seq_documentation_block:
LOCUS       AX031279      332 bp    DNA             PAT       20-SEP-2000
DEFINITION  Sequence 1 from Patent WO9914321.
ACCESSION   AX031279
VERSION     AX031279.1  GI:10278610
KEYWORDS    .
SOURCE      unidentified.
  ORGANISM  unidentified
            unclassified.
REFERENCE   1  (bases 1 to 332)
  AUTHORS   O'Reilly,L., Puthalakath,H., Adams,J., O'Connor,L., Cory,S.,
            Huang,D.C. and Strasser,A.
  TITLE     Novel therapeutic molecules
  JOURNAL   Patent: WO 9914321-A 1 25-MAR-1999;
            INST MEDICAL W & E HALL (AU) ; PUTHALAKATH HAMSA (AU) ; REILLY
            LORRAINE O (AU) ; ADAMS JERRY (AU) ; CONNOR LIAM O (AU) ; CORY
            SUZANNE (AU) ; HUANG DAVID C S (AU) ; STRASSER ANDREAS (AU)
FEATURES             Location/Qualifiers
     source          1. .332
                     /organism="unidentified"
                     /db_xref="taxon:32644"
     CDS             1. .>330
                     /note="unnamed protein product"
                     /codon_start=1
                     /protein_id="CAC09653.1"
                     /db_xref="GI:10278611"
                     /translation="MAKQPSDVSSECDREGGQLQPAERPPQLRPGAPTSLQTEPQASI
                     RQSQEEPEDLRPEIRIAQELRRIGDEFNETYTRRVFANDYREAEDHPQMVILQLLRFI
                     FRLVWRRH"
BASE COUNT       87 a     85 c     91 g     69 t
ORIGIN


alignment_scores:

                 Quality:  549.00              Length:    140
                   Ratio:  4.991                 Gaps:      1
       Percent Similarity:  78.571   Percent Identity: 78.571

alignment_block:
 US-09-508-832-4 x AX031279   ..

 Align seg 1/1  to: AX031279  from: 1  to: 332

       1 MetAlaLysGlnProSerAspValSerSerGluCysAspArgGluGlyGl 17
         ||||||||||||||||||||||||||||||||||||||||||||||||||
       1 ATGGCCAAGCAACCTTCTGATGTAAGTTCTGAGTGTGACAGAGAAGGTGG 50

      17 yGlnLeuGlnProAlaGluArgProProGlnLeuArgProGlyAlaProT 34
         ||||||||||||||||||||||||||||||||||||||||||||||||||
      51 ACAATTGCAGCCTGCTGAGAGGCCTCCCCAGCTCAGGCCTGGGGCCCCTA 100

      34 hrSerLeuGlnThrGluProGlnAspArgSerProAlaProMetSerCys 50
         ||||||||||||||||||||
     101 CCTCCCTACAGACAGAACCGCAA..........................  123

      51 AspLysSerThrGlnThrProSerProProCysGlnAlaPheAsnHisTy 67

     123 ..................................................  123

      67 rLeuSerAlaMetAlaSerIleArgGlnSerGlnGluGluProGluAspL 84
                    ||||||||||||||||||||||||||||||||||||
     124 .............GCTTCCATACGACAGTCTCAGGAGGAACCTGAAGATC 160

      84 euArgProGluIleArgIleAlaGlnGluLeuArgArgIleGlyAspGlu 100
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     161 TGCGCCCGGAGATACGGATTGCACAGGAGCTGCGGCGGATCGGAGACGAG 210

     101 PheAsnGluThrTyrThrArgArgValPheAlaAsnAspTyrArgGluAl 117
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     211 TTCAACGAAACTTACACAAGGAGGGTGTTTGCAAATGATTACCGCGAGGC 260

     117 aGluAspHisProGlnMetValIleLeuGlnLeuLeuArgPheIlePheA 134
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     261 TGAAGACCACCCTCAAATGGTTATCTTACAACTGTTACGCTTTATCTTCC 310

     134 rgLeuValTrpArgArgHis 140
         ||||||||||||||||||
     311 GTCTGGTATGGAGAAGGCAT 330
```

OM of: US-09-508-832-4 to: N_Geneseq_1101:* out_format : pfs

Date: Dec 11, 2001 1:49 AM

About: Results were produced by the GenCore software, version 4.5,
Copyright (c) 1993-2000 Compugen Ltd.

Command line parameters:
-MODEL=frame+_p2n.model -DEV=xlp
-Q=/cgn2_1/USPTO_spool/US09508832/runat_10122001_110349_29569/app_query.fasta_1.620
-DB=N_Geneseq_1101 -QFMT=fastap -SUFFIX=rng -GAPOP=12.000
-GAPEXT=4.000 -MINMATCH=0.100 -LOOPCL=0.000 -LOOPEXT=0.000
-QGAPOP=4.500 -QGAPEXT=0.050 -XGAPOP=10.000 -XGAPEXT=0.500
-FGAPOP=6.000 -FGAPEXT=7.000 -YGAPOP=10.000 -YGAPEXT=0.500
-DELOP=6.000 -DELEXT=7.000 -START=1 -MATRIX=blosum62
-TRANS=human40.cdi -LIST=45 -DOCALIGN=200 -THR_SCORE=pct
-THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=pfs
-NORM=ext -MINLEN=0 -MAXLEN=2000000000
-USER=US09508832_@CGN1_1_327 -NCPU=6 -ICPU=3 -LONGLOG -NO_XLPXY
-WAIT -THREADS=1

Search information block:
Query: US-09-508-832-4
Query length: 140
Database: N_Geneseq_1101:*
Database sequences: 930621
Database length: 428662619
Search time (sec): 249.590000

score_list:

Sequence	Strd	Orig	ZScore	EScore	Len	Documentation
/SIDS2/gcgdata/geneseq/geneseqn/NA1999.DAT:AAX24994	+	742.00	1205.41	4.6e-59	422	
/SIDS2/gcgdata/geneseq/geneseqn/NA1999.DAT:AAX24995	+	704.00	1140.55	1.9e-55	590	
/SIDS2/gcgdata/geneseq/geneseqn/NA1999.DAT:AAX24996	+	629.00	1021.84	7.7e-49	416	
/SIDS2/gcgdata/geneseq/geneseqn/NA1999.DAT:AAX24997	+	589.00	953.50	4.9e-45	596	
/SIDS2/gcgdata/geneseq/geneseqn/NA1999.DAT:AAX24993	+	549.00	893.85	1.0e-41	332	
/SIDS2/gcgdata/geneseq/geneseqn/NA2000.DAT:AAC55832	+	92.50	141.06	8.83	1056	
/SIDS2/gcgdata/geneseq/geneseqn/NA2000.DAT:AAC55842	+	92.50	104.94	907.52	53500	
/SIDS2/gcgdata/geneseq/geneseqn/NA2001.DAT:AAS23095	+	91.50	126.19	59.42	4452	
/SIDS2/gcgdata/geneseq/geneseqn/NA2001.DAT:AAF90327	-	90.50	125.48	65.13	4033	
/SIDS2/gcgdata/geneseq/geneseqn/NA1999.DAT:AAZ32020	-	90.50	104.66	940.62	38734	
/SIDS2/gcgdata/geneseq/geneseqn/NA2001.DAT:AAC90077	-	90.50	104.66	940.62	38734	
/SIDS2/gcgdata/geneseq/geneseqn/NA2001.DAT:AAI21994	-	89.00	124.80	71.06	3331	
/SIDS2/gcgdata/geneseq/geneseqn/NA2001.DAT:AAI47284	-	89.00	124.80	71.06	3331	
/SIDS2/gcgdata/geneseq/geneseqn/NA2001.DAT:AAI07689	-	89.00	124.80	71.06	3331	
/SIDS2/gcgdata/geneseq/geneseqn/NA1999.DAT:AAZ32025	-	89.00	114.86	254.27	9810	
/SIDS2/gcgdata/geneseq/geneseqn/NA2001.DAT:AAC90082	-	89.00	114.86	254.27	9810	
/SIDS2/gcgdata/geneseq/geneseqn/NA2001.DAT:AAI12815	-	88.50	128.73	42.89	1988	
/SIDS2/gcgdata/geneseq/geneseqn/NA2001.DAT:AAI34170	-	88.50	128.73	42.89	1988	
/SIDS2/gcgdata/geneseq/geneseqn/NA2001.DAT:AAI02732	-	88.50	128.73	42.89	1988	
/SIDS2/gcgdata/geneseq/geneseqn/NA1995.DAT:AAQ92807	+	88.50	124.03	78.39	3314	
/SIDS2/gcgdata/geneseq/geneseqn/NA2000.DAT:AAA72948	+	88.50	124.03	78.39	3314	
/SIDS2/gcgdata/geneseq/geneseqn/NA2001.DAT:AAI57831	-	88.50	121.90	102.99	4176	
/SIDS2/gcgdata/geneseq/geneseqn/NA2001.DAT:AAS23001	+	88.50	121.33	110.83	4444	
/SIDS2/gcgdata/geneseq/geneseqn/NA2000.DAT:AAA59847	-	88.00	138.18	12.77	652	
/SIDS2/gcgdata/geneseq/geneseqn/NA2000.DAT:AAC99089	-	88.00	129.16	40.62	1738	
/SIDS2/gcgdata/geneseq/geneseqn/NA2000.DAT:AAA59875	-	88.00	128.18	46.05	1933	
/SIDS2/gcgdata/geneseq/geneseqn/NA2000.DAT:AAZ35835	-	88.00	128.18	46.05	1933	
/SIDS2/gcgdata/geneseq/geneseqn/NA1996.DAT:AAT44385	+	88.00	107.19	679.68	18912	
/SIDS2/gcgdata/geneseq/geneseqn/NA1996.DAT:AAT44385	-	88.00	107.19	679.68	18912	
/SIDS2/gcgdata/geneseq/geneseqn/NA2000.DAT:AAA97430	+	87.50	120.41	124.68	4115	
/SIDS2/gcgdata/geneseq/geneseqn/NA2000.DAT:AAF20955	+	87.50	115.70	228.30	6870	
/SIDS2/gcgdata/geneseq/geneseqn/NA2000.DAT:AAA34833	+	87.50	115.70	228.30	6870	
/SIDS2/gcgdata/geneseq/geneseqn/NA2000.DAT:AAF20956	+	87.50	114.23	275.47	8055	
/SIDS2/gcgdata/geneseq/geneseqn/NA2000.DAT:AAA34834	+	87.50	114.23	275.47	8055	
/SIDS2/gcgdata/geneseq/geneseqn/NA2000.DAT:AAF21437	+	87.50	84.26	1.3e+04	209273	
/SIDS2/gcgdata/geneseq/geneseqn/NA1995.DAT:AAQ86927	-	86.50	117.09	190.87	4947	
/SIDS2/gcgdata/geneseq/geneseqn/NA2000.DAT:AAA59701	+	86.00	116.19	214.42	4998	
/SIDS2/gcgdata/geneseq/geneseqn/NA2000.DAT:AAF12769	+	85.00	128.99	41.51	1042	
/SIDS2/gcgdata/geneseq/geneseqn/NA2000.DAT:AAC41059	+	84.50	135.58	17.82	466	
/SIDS2/gcgdata/geneseq/geneseqn/NA1997.DAT:AAT99437	+	84.50	125.85	62.10	1342	
/SIDS2/gcgdata/geneseq/geneseqn/NA1998.DAT:AAV58307	+	84.50	125.83	62.27	1345	
/SIDS2/gcgdata/geneseq/geneseqn/NA1998.DAT:AAT86631	+	84.50	125.83	62.27	1345	
/SIDS2/gcgdata/geneseq/geneseqn/NA1998.DAT:AAV02763	+	84.50	125.83	62.27	1345	
/SIDS2/gcgdata/geneseq/geneseqn/NA1998.DAT:AAV06021	+	84.50	125.83	62.27	1345	
/SIDS2/gcgdata/geneseq/geneseqn/NA2000.DAT:AAC61410	+	84.50	125.83	62.27	1345	

seq_name: /SIDS2/gcgdata/geneseq/geneseqn/NA1999.DAT:AAX24994

seq_documentation_block:
ID AAX24994 standard; cDNA; 422 BP.
XX
AC AAX24994;
XX
DT 05-JUL-1999 (first entry)
XX
DE Murine Bcl-2 interacting mediator of cell death Bim-L cDNA.
XX
KW Bim-L; Bcl-2 interacting mediator of cell death; apoptosis;
KW cell cycle; mouse; cancer; autoimmune disease;
KW degenerative disease; therapy; contraceptive; splice variant;
KW isoform; ss.
XX
OS Mus musculus.
XX
PN WO9914321-A1.
XX
PD 25-MAR-1999.
XX
PF 17-SEP-1998; 98WO-AU00772.
XX
PR 24-SEP-1997; 97AU-0009373.
PR 17-SEP-1997; 97AU-0009263.
XX
PA (HALL-) HALL INST MEDICAL RES WALTER & ELIZA.
XX
PI Adams J, Cory S, Huang DCS, O'Connor L, O'Reilly L;
PI Puthalakath H, Strasser A;
XX
DR WPI; 1999-244030/20.
DR P-PSDB; AAW98155.
XX
PT New isolated member of the Bcl-2 family, Bim used in, e.g. cancer
PT treatment
XX
PS Claim 3; Page 94-95; 145pp; English.
XX
CC The present sequence encodes the long form (L) of murine Bim, or
CC Bcl-2 interacting mediator of cell death (see AAW98155), a novel
CC member of the Bcl-2 family that is capable of inducing cell death
CC (apoptosis) and which acts as a 'death-ligand' for certain members
CC of the pro-survival Bcl-2 family. Bim is a BH3-only protein, as the
CC only Bcl-2 homology region which it encompasses is BH3. It is the
CC only BH3-only protein for which splice variants exist. These
CC result in the expression of a variety of isoforms, i.e. Bim-S,
CC Bim-L and Bim-EL (see AAW98154-56). cDNAs encoding these murine Bim
CC isoforms were obtained from a T lymphoma cDNA library using human
CC recombinant Bcl-2 protein. The murine Bim gene has been mapped to
CC chromosome 2 at bands F3-G. Human Bim-L and Bim-EL isoforms have
CC also been identified (see AAW98157-58). Binding the dynein light
CC chain was shown to regulate the pro-apoptotic activity of Bim.
CC Bim-S, the splice variant which does not bind to dynein light
CC chain, is a much more potent killer than either Bim-L or Bim-EL.
CC The invention provides variants (see AAW98159-68) of murine and human
CC Bim-L or Bim-EL that cannot bind, couple or otherwise associate
CC with a dynein light chain. The identification of Bim permits the
CC identification and rational design of a range of products for use
CC in therapy, diagnosis, antibody generation and involving modulation
CC of physiological cell death. These therapeutic molecules may act
CC as either antagonists or agonists of Bim's function and will be
CC useful in cancer, autoimmune or degenerative disease therapy.
CC Increased Bim expression or Bim activity is useful, e.g. for
CC treatment or prophylaxis in conditions such as cancer and deletion
CC of autoreactive lymphocytes in autoimmune disease. Decreased Bim

```
CC   expression of Bim activity is useful in regulating inhibition or
CC   prevention of cell death or degeneration such as under cytotoxic
CC   conditions during e.g. gamma-irradiation and chemotherapy or during
CC   HIV/AIDS or other viral infections, ischemia, myocardial infarction,
CC   hypoxia, degenerative diseases or for prolonging the survival of
CC   cells being transplanted for treatment of disease.  Since Bim is
CC   expressed in germ cells, modulating Bim expression or Bim activity
CC   is useful, e.g. as a contraceptive or method of sterilization by
CC   preventing generation of fertile sperm.
XX
SQ   Sequence 422 BP; 112 A; 116 C; 109 G; 85 T; 0 other;


alignment_scores:
            Quality:  742.00              Length:    140
              Ratio:  5.300                Gaps:      0
 Percent Similarity: 100.000   Percent Identity: 100.000

alignment_block:
 US-09-508-832-4 x AAX24994   ..

 Align seg 1/1  to: AAX24994  from: 1  to: 422

       1 MetAlaLysGlnProSerAspValSerSerGluCysAspArgGluGlyGl 17
         ||||||||||||||||||||||||||||||||||||||||||||||||||
       1 ATGGCCAAGCAACCTTCTGATGTAAGTTCTGAGTGTGACAGAGAAGGTGG 50

      17 yGlnLeuGlnProAlaGluArgProProGlnLeuArgProGlyAlaProT 34
         ||||||||||||||||||||||||||||||||||||||||||||||||||
      51 ACAATTGCAGCCTGCTGAGAGGCCTCCCCAGCTCAGGCCTGGGGCCCCTA 100

      34 hrSerLeuGlnThrGluProGlnAspArgSerProAlaProMetSerCys 50
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     101 CCTCCCTACAGACAGAACCGCAAGACAGGAGCCCGGCACCCATGAGTTGT 150

      51 AspLysSerThrGlnThrProSerProProCysGlnAlaPheAsnHisTy 67
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     151 GACAAGTCAACACAAACCCCAAGTCCTCCTTGCCAGGCCTTCAACCACTA 200

      67 rLeuSerAlaMetAlaSerIleArgGlnSerGlnGluGluProGluAspL 84
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     201 TCTCAGTGCAATGGCTTCCATACGACAGTCTCAGGAGGAACCTGAAGATC 250

      84 euArgProGluIleArgIleAlaGlnGluLeuArgArgIleGlyAspGlu 100
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     251 TGCGCCCGGAGATACGGATTGCACAGGAGCTGCGGCGGATCGGAGACGAG 300

     101 PheAsnGluThrTyrThrArgArgValPheAlaAsnAspTyrArgGluAl 117
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     301 TTCAACGAAACTTACACAAGGAGGGTGTTTGCAAATGATTACCGCGAGGC 350

     117 aGluAspHisProGlnMetValIleLeuGlnLeuLeuArgPheIlePheA 134
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     351 TGAAGACCACCCTCAAATGGTTATCTTACAACTGTTACGCTTTATCTTCC 400

     134 rgLeuValTrpArgArgHis 140
         ||||||||||||||||||
     401 GTCTGGTATGGAGAAGGCAT 420

seq_name: /SIDS2/gcgdata/geneseq/geneseqn/NA1999.DAT:AAX24995

seq_documentation_block:
ID   AAX24995 standard; cDNA; 590 BP.
XX
AC   AAX24995;
XX
DT   05-JUL-1999  (first entry)
XX
DE   Murine Bcl-2 interacting mediator of cell death Bim-EL cDNA.
XX
KW   Bim-EL; Bcl-2 interacting mediator of cell death; apoptosis;
KW   cell cycle; mouse; cancer; autoimmune disease;
KW   degenerative disease; therapy; contraceptive; splice variant;
KW   isoform; ss.
XX
OS   Mus musculus.
XX
PN   WO9914321-A1.
XX
PD   25-MAR-1999.
XX
PF   17-SEP-1998;   98WO-AU00772.
XX
PR   24-SEP-1997;   97AU-0009373.
PR   17-SEP-1997;   97AU-0009263.
XX
PA   (HALL-) HALL INST MEDICAL RES WALTER & ELIZA.
XX
PI   Adams J,  Cory S,  Huang DCS,  O'Connor L,  O'Reilly L;
PI   Puthalakath H,  Strasser A;
XX
DR   WPI; 1999-244030/20.
DR   P-PSDB; AAW98156.
XX
PT   New isolated member of the Bcl-2 family, Bim used in, e.g. cancer
PT   treatment
XX
PS   Claim 3; Page 96-97; 145pp; English.
XX
CC   The present sequence encodes the extra long form (EL) of murine Bim,
CC   or Bcl-2 interacting mediator of cell death (see AAW98156), a novel
CC   member of the Bcl-2 family that is capable of inducing cell death
CC   (apoptosis) and which acts as a 'death-ligand' for certain members
CC   of the pro-survival Bcl-2 family.  Bim is a BH3-only protein, as the
CC   only Bcl-2 homology region which it encompasses is BH3.  It is the
CC   only BH3-only protein for which splice variants exist.  These
CC   result in the expression of a variety of isoforms, i.e. Bim-S,
CC   Bim-L and Bim-EL (see AAW98154-56).  cDNAs encoding these murine Bim
CC   isoforms were obtained from a T lymphoma cDNA library using human
CC   recombinant Bcl-2 protein.  The murine Bim gene has been mapped to
CC   chromosome 2 at bands F3-G.  Human Bim-L and Bim-EL isoforms have
CC   also been identified (see AAW98157-58).  Binding the dynein light
CC   chain was shown to regulate the pro-apoptotic activity of Bim.
CC   Bim-S, the splice variant which does not bind to dynein light
CC   chain, is a much more potent killer than either Bim-L or Bim-EL.
CC   The invention provides variants (see AAW98159-68) of murine and human
CC   Bim-L or Bim-EL that cannot bind, couple or otherwise associate
CC   with a dynein light chain.  The identification of Bim permits the
CC   identification and rational design of a range of products for use
CC   in therapy, diagnosis, antibody generation and involving modulation
CC   of physiological cell death.  These therapeutic molecules may act
CC   as either antagonists or agonists of Bim's function and will be
CC   useful in cancer, autoimmune or degenerative disease therapy.
CC   Increased Bim expression or Bim activity is useful, e.g. for
CC   treatment or prophylaxis in conditions such as cancer and deletion
CC   of autoreactive lymphocytes in autoimmune disease.  Decreased Bim
CC   expression of Bim activity is useful in regulating inhibition or
CC   prevention of cell death or degeneration such as under cytotoxic
CC   conditions during e.g. gamma-irradiation and chemotherapy or during
CC   HIV/AIDS or other viral infections, ischemia, myocardial infarction,
CC   hypoxia, degenerative diseases or for prolonging the survival of
CC   cells being transplanted for treatment of disease.  Since Bim is
CC   expressed in germ cells, modulating Bim expression or Bim activity
CC   is useful, e.g. as a contraceptive or method of sterilization by
CC   preventing generation of fertile sperm.
XX
SQ   Sequence 590 BP; 137 A; 178 C; 150 G; 125 T; 0 other;


alignment_scores:
            Quality:  704.00              Length:    196
              Ratio:  5.029                Gaps:      1
 Percent Similarity:  71.429   Percent Identity: 71.429

alignment_block:
 US-09-508-832-4 x AAX24995   ..
```

```
Align seg 1/1  to: AAX24995  from: 1  to: 590

      1 MetAlaLysGlnProSerAspValSerSerGluCysAspArgGluGlyGl 17
        ||||||||||||||||||||||||||||||||||||||||||||||||||
      1 ATGGCCAAGCAACCTTCTGATGTAAGTTCTGAGTGTGACAGAGAAGGTGG 50

     17 yGlnLeuGlnProAlaGluArgProProGlnLeuArgProGlyAlaProT 34
        ||||||||||||||||||||||||||||||||||||||||||||||||||
     51 ACAATTGCAGCCTGCTGAGAGGCCTCCCCAGCTCAGGCCTGGGGCCCCTA 100

     34 hrSerLeuGlnThrGluProGln..........................  41
        ||||||||||||||||||||||
    101 CCTCCCTACAGACAGAACCGCAAGGTAATCCCGACGGCGAAGGGGACCGC 150

     41 .....................................................  41

    151 TGCCCCCACGGCAGCCCTCAGGGCCCGCTGGCCCCACCGGCCAGCCCTGG 200

     41 .....................................................  41

    201 CCCTTTTGCTACCAGATCCCCACTTTTCATCTTTGTGAGAAGATCTTCTC 250

     42 ..................................AspArgSer 44
                                           |||||||||
    251 TGCTGTCCCGGTCCTCCAGTGGGTATTTCTCTTTTGACACAGACAGGAGC 300

     45 ProAlaProMetSerCysAspLysSerThrGlnThrProSerProProCy 61
        ||||||||||||||||||||||||||||||||||||||||||||||||||
    301 CCGGCACCCATGAGTTGTGACAAGTCAACACAAACCCCAAGTCCTCCTTG 350

     61 sGlnAlaPheAsnHisTyrLeuSerAlaMetAlaSerIleArgGlnSerG 78
        ||||||||||||||||||||||||||||||||||||||||||||||||||
    351 CCAGGCCTTCAACCACTATCTCAGTGCAATGGCTTCCATACGACAGTCTC 400

     78 lnGluGluProGluAspLeuArgProGluIleArgIleAlaGlnGluLeu 94
        ||||||||||||||||||||||||||||||||||||||||||||||||||
    401 AGGAGGAACCTGAAGATCTGCGCCCGGAGATACGGATTGCACAGGAGCTG 450

     95 ArgArgIleGlyAspGluPheAsnGluThrTyrThrArgArgValPheAl 111
        ||||||||||||||||||||||||||||||||||||||||||||||||||
    451 CGGCGGATCGGAGACGAGTTCAACGAAACTTACACAAGGAGGGTGTTTGC 500

    111 aAsnAspTyrArgGluAlaGluAspHisProGlnMetValIleLeuGlnL 128
        ||||||||||||||||||||||||||||||||||||||||||||||||||
    501 AAATGATTACCGCGAGGCTGAAGACCACCCTCAAATGGTTATCTTACAAC 550

    128 euLeuArgPheIlePheArgLeuValTrpArgArgHis 140
        |||||||||||||||||||||||||||||||||||||
    551 TGTTACGCTTTATCTTCCGTCTGGTATGGAGAAGGCAT 588

seq_name: /SIDS2/gcgdata/geneseq/geneseqn/NA1999.DAT:AAX24996

seq_documentation_block:
ID   AAX24996 standard; cDNA; 416 BP.
XX
AC   AAX24996;
XX
DT   05-JUL-1999  (first entry)
XX
DE   Human Bcl-2 interacting mediator of cell death Bim-L cDNA.
XX
KW   Bim-L; Bcl-2 interacting mediator of cell death; apoptosis;
KW   cell cycle; human; cancer; autoimmune disease;
KW   degenerative disease; therapy; contraceptive; splice variant;
KW   isoform; ss.
XX
OS   Homo sapiens.
XX
PN   WO9914321-A1.
XX
PD   25-MAR-1999.
XX
PF   17-SEP-1998;   98WO-AU00772.
XX
PR   24-SEP-1997;   97AU-0009373.
PR   17-SEP-1997;   97AU-0009263.
XX
PA   (HALL-) HALL INST MEDICAL RES WALTER & ELIZA.
XX
PI   Adams J,  Cory S,  Huang DCS,  O'Connor L,  O'Reilly L;
PI   Puthalakath H,  Strasser A;
XX
DR   WPI; 1999-244030/20.
DR   P-PSDB; AAW98157.
XX
PT   New isolated member of the Bcl-2 family, Bim used in, e.g. cancer
PT   treatment
XX
PS   Claim 7; Page 99-100; 145pp; English.
XX
CC   The present sequence encodes the long form (L) of human Bim, or
CC   Bcl-2 interacting mediator of cell death (see AAW98157), a novel
CC   member of the Bcl-2 family that is capable of inducing cell death
CC   (apoptosis) and which acts as a 'death-ligand' for certain members
CC   of the pro-survival Bcl-2 family.  Bim is a BH3-only protein, as the
CC   only Bcl-2 homology region which it encompasses is BH3.  It is the
CC   only BH3-only protein for which splice variants exist.  These
CC   result in the expression of a variety of isoforms, i.e. Bim-S,
CC   Bim-L and Bim-EL.  cDNAs encoding human Bim-L and Bim-EL (see
CC   AAW98158) were isolated from embryo and liver cDNA libraries using
CC   mouse bim cDNA.  Murine Bim-S, Bim-L and Bim-EL isoforms (see
CC   AAW98154-56) are also provided.  The human Bim gene maps to
CC   chromosome 2 at bands 2q12-2q13.  Binding the dynein light
CC   chain was shown to regulate the pro-apoptotic activity of Bim.
CC   Bim-S, the splice variant which does not bind to dynein light
CC   chain, is a much more potent killer than either Bim-L or Bim-EL.
CC   The invention provides variants (see AAW98159-68) of murine and human
CC   Bim-L or Bim-EL that cannot bind, couple or otherwise associate
CC   with a dynein light chain.  The identification of Bim permits the
CC   identification and rational design of a range of products for use
CC   in therapy, diagnosis, antibody generation and involving modulation
CC   of physiological cell death.  These therapeutic molecules may act
CC   as either antagonists or agonists of Bim's function and will be
CC   useful in cancer, autoimmune or degenerative disease therapy.
CC   Increased Bim expression or Bim activity is useful, e.g. for
CC   treatment or prophylaxis in conditions such as cancer and deletion
CC   of autoreactive lymphocytes in autoimmune disease.  Decreased Bim
CC   expression of Bim activity is useful in regulating inhibition or
CC   prevention of cell death or degeneration such as under cytotoxic
CC   conditions during e.g. gamma-irradiation and chemotherapy or during
CC   HIV/AIDS or other viral infections, ischemia, myocardial infarction,
CC   hypoxia, degenerative diseases or for prolonging the survival of
CC   cells being transplanted for treatment of disease.  Since Bim is
CC   expressed in germ cells, modulating Bim expression or Bim activity
CC   is useful, e.g. as a contraceptive or method of sterilization by
CC   preventing generation of fertile sperm.
XX
SQ   Sequence 416 BP; 113 A; 113 C; 103 G; 87 T; 0 other;

alignment_scores:
          Quality:  629.00              Length:    140
            Ratio:  4.876                 Gaps:      1
 Percent Similarity:  92.143   Percent Identity: 85.714

alignment_block:
 US-09-508-832-4 x AAX24996   ..

 Align seg 1/1  to: AAX24996  from: 1  to: 416

       1 MetAlaLysGlnProSerAspValSerSerGluCysAspArgGluGlyGl 17
         ||||||||||||||||||||||||||||||||||||||||||||||||||
       1 ATGGCAAAGCAACCTTCTGATGTAAGTTCTGAGTGTGACCGAGAAGGTAG 50

      17 yGlnLeuGlnProAlaGluArgProProGlnLeuArgProGlyAlaProT 34
```

```
      ||||||||||||||||||||||||||||||||||||||||||||||||||
   51 ACAATTGCAGCCTGCGGAGAGGCCTCCCCAGCTCAGACCTGGGGCCCCTA 100

   34 hrSerLeuGlnThrGluProGlnAspArgSerProAlaProMetSerCys 50
      ||||||||||||||||||||||||||||||||||||||||||||||||||
  101 CCTCCCTACAGACAGAGCCACAAGACAGGAGCCCAGCACCCATGAGTTGT 150

   51 AspLysSerThrGlnThrProSerProProCysGlnAlaPheAsnHisTy 67
      ||||||||||||||||||||||||||||||||||||||||||||||||||
  151 GACAAATCAACACAAACCCCAAGTCCTCCTTGCCAGGCCTTCAACCACTA 200

   67 rLeuSerAlaMetAlaSerIleArgGlnSerGlnGluGluProGluAspL 84
      ||||||||||||||||||:::||||||:::      ||||||   |||:
  201 TCTCAGTGCAATGGCTTCCATGAGGCAGGCT......GAACCTGCAGATA 244

   84 euArgProGluIleArgIleAlaGlnGluLeuArgArgIleGlyAspGlu 100
      ::|||||||||||||   ||||||||||||||||||||||||||||||||
  245 TGCGCCCAGAGATATGGATCGCCCAAGAGTTGCGGCGTATCGGAGACGAG 294

  101 PheAsnGluThrTyrThrArgArgValPheAlaAsnAspTyrArgGluAl 117
      ||||||      |||:::||||||||||||   |||:::|||:::   ||
  295 TTTAACGCTTACTATGCAAGGAGGGTATTTTTGAATAATTACCAAGCAGC 344

  117 aGluAspHisProGlnMetValIleLeuGlnLeuLeuArgPheIlePheA 134
      ||||||||||||:::||||||||||||:::||||||||||:::|||   |
  345 CGAAGACCACCCACGAATGGTTATCTTACGACTGTTACGTTACATTGTCC 394

  134 rgLeuValTrpArgArgHis 140
      ||||||||||||||   |||
  395 GCCTGGTGTGGAGAATGCAT 414

seq_name: /SIDS2/gcgdata/geneseq/geneseqn/NA1999.DAT:AAX24997

seq_documentation_block:
ID   AAX24997 standard; cDNA; 596 BP.
XX
AC   AAX24997;
XX
DT   05-JUL-1999  (first entry)
XX
DE   Human Bcl-2 interacting mediator of cell death Bim-EL cDNA.
XX
KW   Bim-EL; Bcl-2 interacting mediator of cell death; apoptosis;
KW   cell cycle; human; cancer; autoimmune disease;
KW   degenerative disease; therapy; contraceptive; splice variant;
KW   isoform; ss.
XX
OS   Homo sapiens.
XX
PN   WO9914321-A1.
XX
PD   25-MAR-1999.
XX
PF   17-SEP-1998;   98WO-AU00772.
XX
PR   24-SEP-1997;   97AU-0009373.
PR   17-SEP-1997;   97AU-0009263.
XX
PA   (HALL-) HALL INST MEDICAL RES WALTER & ELIZA.
XX
PI   Adams J,  Cory S,  Huang DCS,  O'Connor L,  O'Reilly L;
PI   Puthalakath H,  Strasser A;
XX
DR   WPI; 1999-244030/20.
DR   P-PSDB; AAW98158.
XX
PT   New isolated member of the Bcl-2 family, Bim used in, e.g. cancer
PT   treatment
XX
PS   Claim 7; Page 101-102; 145pp; English.
XX
CC   The present sequence encodes the extra long form (EL) of human Bim,
CC   or Bcl-2 interacting mediator of cell death (see AAW98158), a novel
CC   member of the Bcl-2 family that is capable of inducing cell death
CC   (apoptosis) and which acts as a 'death-ligand' for certain members
CC   of the pro-survival Bcl-2 family.  Bim is a BH3-only protein, as the
CC   only Bcl-2 homology region which it encompasses is BH3.  It is the
CC   only BH3-only protein for which splice variants exist.  These
CC   result in the expression of a variety of isoforms, i.e. Bim-S,
CC   Bim-L and Bim-EL.  cDNAs encoding human Bim-L and Bim-EL (see
CC   AAW98158) were isolated from embryo and liver cDNA libraries using
CC   mouse bim cDNA.  Murine Bim-S, Bim-L and Bim-EL isoforms (see
CC   AAW98154-56) are also provided.  The human Bim gene maps to
CC   chromosome 2 at bands 2q12-2q13.  Binding the dynein light
CC   chain was shown to regulate the pro-apoptotic activity of Bim.
CC   Bim-S, the splice variant which does not bind to dynein light
CC   chain, is a much more potent killer than either Bim-L or Bim-EL.
CC   The invention provides variants (see AAW98159-68) of murine and human
CC   Bim-L or Bim-EL that cannot bind, couple or otherwise associate
CC   with a dynein light chain.  The identification of Bim permits the
CC   identification and rational design of a range of products for use
CC   in therapy, diagnosis, antibody generation and involving modulation
CC   of physiological cell death.  These therapeutic molecules may act
CC   as either antagonists or agonists of Bim's function and will be
CC   useful in cancer, autoimmune or degenerative disease therapy.
CC   Increased Bim expression or Bim activity is useful, e.g. for
CC   treatment or prophylaxis in conditions such as cancer and deletion
CC   of autoreactive lymphocytes in autoimmune disease.  Decreased Bim
CC   expression of Bim activity is useful in regulating inhibition or
CC   prevention of cell death or degeneration such as under cytotoxic
CC   conditions during e.g. gamma-irradiation and chemotherapy or during
CC   HIV/AIDS or other viral infections, ischemia, myocardial infarction,
CC   hypoxia, degenerative diseases or for prolonging the survival of
CC   cells being transplanted for treatment of disease.  Since Bim is
CC   expressed in germ cells, modulating Bim expression or Bim activity
CC   is useful, e.g. as a contraceptive or method of sterilization by
CC   preventing generation of fertile sperm.
XX
SQ   Sequence 596 BP; 145 A; 175 C; 146 G; 130 T; 0 other;


alignment_scores:
           Quality:  589.00              Length:    200
             Ratio:  4.566                 Gaps:      2
 Percent Similarity:  64.500   Percent Identity: 60.000

alignment_block:
 US-09-508-832-4 x AAX24997   ..

 Align seg 1/1  to: AAX24997  from: 1  to: 596

      1 MetAlaLysGlnProSerAspValSerSerGluCysAspArgGluGlyGl 17
        ||||||||||||||||||||||||||||||||||||||||||||||||||
      1 ATGGCAAAGCAACCTTCTGATGTAAGTTCTGAGTGTGACCGAGAAGGTAG 50

     17 yGlnLeuGlnProAlaGluArgProProGlnLeuArgProGlyAlaProT 34
        ||||||||||||||||||||||||||||||||||||||||||||||||||
     51 ACAATTGCAGCCTGCGGAGAGGCCTCCCCAGCTCAGACCTGGGGCCCCTA 100

     34 hrSerLeuGlnThrGluProGln..................... 41
        ||||||||||||||||||||||
    101 CCTCCCTACAGACAGAGCCACAAGGTAATCCTGAAGGCAATCACGGAGGT 150

     41 ..................................................... 41

    151 GAAGGGGACAGCTGCCCCCACGGCAGCCCTCAGGGCCCGCTGGCCCCACC 200

     41 ..................................................... 41

    201 TGCCAGCCCTGGCCCTTTTGCTACCAGATCCCCGCTTTTCATCTTTATGA 250

     41 ..................................................... 41

    251 GAAGATCCTCCCTGCTGTCTCGATCCTCCAGTGGGTATTTCTCTTTTGAC 300

     42 ...AspArgSerProAlaProMetSerCysAspLysSerThrGlnThrPr 57
```

```
          ||||||||||||||||||||||||||||||||||||||||||||||||
     301 ACAGACAGGAGCCCAGCACCCATGAGTTGTGACAAATCAACACAAACCCC 350

      57 oSerProProCysGlnAlaPheAsnHisTyrLeuSerAlaMetAlaSerI 74
         ||||||||||||||||||||||||||||||||||||||||||||||||:
     351 AAGTCCTCCTTGCCAGGCCTTCAACCACTATCTCAGTGCAATGGCTTCCA 400

      74 leArgGlnSerGlnGluGluProGluAspLeuArgProGluIleArgIle 90
         ::||||||:::      ||||||   |||:::|||||||||||||   |||
     401 TGAGGCAGGCT......GAACCTGCAGATATGCGCCCAGAGATATGGATC 444

      91 AlaGlnGluLeuArgArgIleGlyAspGluPheAsnGluThrTyrThrAr 107
         ||||||||||||||||||||||||||||||||||||      |||:::||
     445 GCCCAAGAGTTGCGGCGTATCGGAGACGAGTTTAACGCTTACTATGCAAG 494

     107 gArgValPheAlaAsnAspTyrArgGluAlaGluAspHisProGlnMetV 124
         ||||||||||   |||:::|||:::   ||||||||||||||||:::|||| 
     495 GAGGGTATTTTTGAATAATTACCAAGCAGCCGAAGACCACCCACGAATGG 544

     124 alIleLeuGlnLeuLeuArgPheIlePheArgLeuValTrpArgArgHis 140
         ||||||||:::|||||||||:::|||   ||||||||||||||   |||
     545 TTATCTTACGACTGTTACGTTACATTGTCCGCCTGGTGTGGAGAATGCAT 594
```

seq_name: /SIDS2/gcgdata/geneseq/geneseqn/NA1999.DAT:AAX24993

seq_documentation_block:

```
ID   AAX24993 standard; cDNA; 332 BP.
XX
AC   AAX24993;
XX
DT   05-JUL-1999  (first entry)
XX
DE   Murine Bcl-2 interacting mediator of cell death Bim-S cDNA.
XX
KW   Bim-S; Bcl-2 interacting mediator of cell death; apoptosis;
KW   cell cycle; mouse; cancer; autoimmune disease;
KW   degenerative disease; therapy; contraceptive; splice variant;
KW   isoform; ss.
XX
OS   Mus musculus.
XX
PN   WO9914321-A1.
XX
PD   25-MAR-1999.
XX
PF   17-SEP-1998;   98WO-AU00772.
XX
PR   24-SEP-1997;   97AU-0009373.
PR   17-SEP-1997;   97AU-0009263.
XX
PA   (HALL-) HALL INST MEDICAL RES WALTER & ELIZA.
XX
PI   Adams J,  Cory S,  Huang DCS,  O'Connor L,  O'Reilly L;
PI   Puthalakath H,  Strasser A;
XX
DR   WPI; 1999-244030/20.
DR   P-PSDB; AAW98154.
XX
PT   New isolated member of the Bcl-2 family, Bim used in, e.g. cancer
PT   treatment
XX
PS   Claim 3; Page 92; 145pp; English.
XX
CC   The present sequence encodes the short form (S) of murine Bim, or
CC   Bcl-2 interacting mediator of cell death (see AAW98154), a novel
CC   member of the Bcl-2 family that is capable of inducing cell death
CC   (apoptosis) and which acts as a 'death-ligand' for certain members
CC   of the pro-survival Bcl-2 family.  Bim is a BH3-only protein, as the
CC   only Bcl-2 homology region which it encompasses is BH3.  It is the
CC   only BH3-only protein for which splice variants exist.  These
CC   result in the expression of a variety of isoforms, i.e. Bim-S,
CC   Bim-L and Bim-EL (see AAW98154-56).  cDNAs encoding these murine Bim
CC   isoforms were obtained from a T lymphoma cDNA library using human
CC   recombinant Bcl-2 protein.  The murine Bim gene has been mapped to
CC   chromosome 2 at bands F3-G.  Human Bim-L and Bim-EL isoforms have
CC   also been identified (see AAW98157-58).  Binding the dynein light
CC   chain was shown to regulate the pro-apoptotic activity of Bim.
CC   Bim-S, the splice variant which does not bind to dynein light
CC   chain, is a much more potent killer than either Bim-L or Bim-EL.
CC   The invention provides variants (see AAW98159-68) of murine and human
CC   Bim-L or Bim-EL that cannot bind, couple or otherwise associate
CC   with a dynein light chain.  The identification of Bim permits the
CC   identification and rational design of a range of products for use
CC   in therapy, diagnosis, antibody generation and involving modulation
CC   of physiological cell death.  These therapeutic molecules may act
CC   as either antagonists or agonists of Bim's function and will be
CC   useful in cancer, autoimmune or degenerative disease therapy.
CC   Increased Bim expression or Bim activity is useful, e.g. for
CC   treatment or prophylaxis in conditions such as cancer and deletion
CC   of autoreactive lymphocytes in autoimmune disease.  Decreased Bim
CC   expression of Bim activity is useful in regulating inhibition or
CC   prevention of cell death or degeneration such as under cytotoxic
CC   conditions during e.g. gamma-irradiation and chemotherapy or during
CC   HIV/AIDS or other viral infections, ischemia, myocardial infarction,
CC   hypoxia, degenerative diseases or for prolonging the survival of
CC   cells being transplanted for treatment of disease.  Since Bim is
CC   expressed in germ cells, modulating Bim expression or Bim activity
CC   is useful, e.g. as a contraceptive or method of sterilization by
CC   preventing generation of fertile sperm.
XX
SQ   Sequence 332 BP; 87 A; 85 C; 91 G; 69 T; 0 other;
```

alignment_scores:

```
            Quality:  549.00              Length:    140
              Ratio:  4.991                 Gaps:      1
 Percent Similarity:  78.571   Percent Identity: 78.571
```

alignment_block:

US-09-508-832-4 x AAX24993 ..

Align seg 1/1 to: AAX24993 from: 1 to: 332

```
          1 MetAlaLysGlnProSerAspValSerSerGluCysAspArgGluGlyGl 17
            ||||||||||||||||||||||||||||||||||||||||||||||||||
          1 ATGGCCAAGCAACCTTCTGATGTAAGTTCTGAGTGTGACAGAGAAGGTGG 50

         17 yGlnLeuGlnProAlaGluArgProProGlnLeuArgProGlyAlaProT 34
            ||||||||||||||||||||||||||||||||||||||||||||||||||
         51 ACAATTGCAGCCTGCTGAGAGGCCTCCCCAGCTCAGGCCTGGGGCCCCTA 100

         34 hrSerLeuGlnThrGluProGlnAspArgSerProAlaProMetSerCys 50
            ||||||||||||||||||||||
        101 CCTCCCTACAGACAGAACCGCAA..................................... 123

         51 AspLysSerThrGlnThrProSerProProCysGlnAlaPheAsnHisTy 67

        123 ..................................................... 123

         67 rLeuSerAlaMetAlaSerIleArgGlnSerGlnGluGluProGluAspL 84
                        ||||||||||||||||||||||||||||||||||||||
        124 ..............GCTTCCATACGACAGTCTCAGGAGGAACCTGAAGATC 160

         84 euArgProGluIleArgIleAlaGlnGluLeuArgArgIleGlyAspGlu 100
            ||||||||||||||||||||||||||||||||||||||||||||||||||
        161 TGCGCCCGGAGATACGGATTGCACAGGAGCTGCGGCGGATCGGAGACGAG 210

        101 PheAsnGluThrTyrThrArgArgValPheAlaAsnAspTyrArgGluAl 117
            ||||||||||||||||||||||||||||||||||||||||||||||||||
        211 TTCAACGAAACTTACACAAGGAGGGTGTTTGCAAATGATTACCGCGAGGC 260

        117 aGluAspHisProGlnMetValIleLeuGlnLeuLeuArgPheIlePheA 134
            ||||||||||||||||||||||||||||||||||||||||||||||||||
        261 TGAAGACCACCCTCAAATGGTTATCTTACAACTGTTACGCTTTATCTTCC 310

        134 rgLeuValTrpArgArgHis 140
```

```
          ||||||||||||||||||||
      311 GTCTGGTATGGAGAAGGCAT 330

seq_name: /SIDS2/gcgdata/geneseq/geneseqn/NA2000.DAT:AAC55832

seq_documentation_block:
ID   AAC55832 standard; DNA; 1056 BP.
XX
AC   AAC55832;
XX
DT   19-JAN-2001  (first entry)
XX
DE   S. lavendulae MmcR encoding DNA sequence.
XX
KW   Mitomycin; biosynthesis; mitosane ring system; antibiotic; anti-cancer;
KW   anti-inflammatory; immune-enhancer; immunosuppressant; asthma;
KW   chronic obstructive pulmonary disease; respiratory inflammation;
KW   fungicide; pesticide; ds.
XX
OS   Streptomyces lavendulae.
XX
PN   WO200053737-A2.
XX
PD   14-SEP-2000.
XX
PF   10-MAR-2000; 2000WO-US06394.
XX
PR   12-MAR-1999;   99US-0266965.
XX
PA   (MINU ) UNIV MINNESOTA.
PA   (SHER/) SHERMAN D H.
PA   (MAOY/) MAO Y.
PA   (VARO/) VAROGLU M.
PA   (HEMM/) HE M.
PA   (SHEL/) SHELDON P C.
XX
PI   Sherman DH,  Mao Y,  Varoglu M,  He M,  Sheldon PC;
XX
DR   WPI; 2000-601980/57.
DR   P-PSDB; AAB32531.
XX
PT   Novel nucleic acid molecule comprising mitomycin biosynthetic gene
PT   cluster useful for cloning mitomycin biosynthetic genes for elucidating
PT   the molecular basis of mitosane ring system biosynthesis  -
XX
PS   Disclosure; Page 281; 399pp; English.
XX
CC   This invention relates to isolated and purified nucleic acid molecules
CC   from the mitomycin biosynthetic gene cluster. Mitomycins are a group of
CC   natural products that contain a variety of functional groups, including
CC   amino benzoquinone and axiridine ring systems. The S. lavendulae
CC   mitomycin biosynthetic gene cluster comprises 47 mitomycin genes
CC   spanning 55kb of DNA. The invention includes an expression cassette
CC   comprising a mitomycin biosynthetic gene operably linked to a promoter,
CC   and host cells transformed with the cassette. The nucleotide, and protein
CC   sequences and the transformed host cells of the invention result in
CC   antiasthmatic, antiinflammatory, cytostatic, immunomodulatory, and
CC   antibiotic activities. The nucleotide sequences are used to elucidate the
CC   molecular basis for the biosynthesis of the mitosane ring system, as well
CC   as to engineer the biosynthesis of novel natural products, e.g.
CC   antibiotics, anti-inflammatory agents, anti-cancer agents,
CC   immune-enhancers, immunosuppressants, agents to treat asthma, chronic
CC   obstructive pulmonary disease as well as other disease involving
CC   respiratory inflammation, or cholesterol-lowering agents or as crop
CC   protection agents (e.g. fungicides or insecticides) as well as
CC   bipolymers, e.g., in packaging or biomedical applications, or to engineer
CC   PHA monomer synthases. Sequences AAC55782-C55881, AAC55815-C55849 and
CC   AAB32485-B32542 represent mitomycin biosynthetic gene cluster DNA
CC   sequences and encoded proteins. Sequences AAC55812-C55814,
CC   AAC55850-C55856 and AAC55862-C55869 represent PCR primers used in the
CC   cloning of the mitomycin biosynthetic genes.
XX
SQ   Sequence 1056 BP; 152 A; 427 C; 338 G; 139 T; 0 other;

alignment_scores:
          Quality:   92.50              Length:    118
            Ratio:  1.779                Gaps:      5
 Percent Similarity:  44.068   Percent Identity: 28.814

alignment_block:
US-09-508-832-4 x AAC55832   ..

Align seg 1/1 to: AAC55832  from: 1  to: 1056

       5 ProSerAsp...ValSerSerGluCysAspArgGluGlyGlyGlnLeuGl 20
         ||||||:::   :::|||   :::|||:::|||::::::
     113 CCGTCGAACTCGGCGTCCCGGAACTGCTCCAGGAGGGCC..... 151

      20 nProAlaGluArgProProGlnLeuArgProGlyAlaPro..... 33
          ||||||   |||||||||:::   ||||||   ||||||
     152 .CCCGCACCGCGACCGCCCTCGCCGAGGCCACCGGCGCCCACGAGCAGAC 200

      34 ..........................................Thr 34
                                                   |||
     201 CCTGCGCAGACTGCTCCGACTGCTCGCCACGGTGGGCGTCTTCGACGACC 250

      35 SerLeuGlnThrGluProGlnAspArgSerProAlaProMetSerCysAs 51
         |||      |||      :::   |||:::|||:::|||   ||||||
     251 TCGGCCACGACGACCTGTTCGCCCAGAACGCCCTCTCCGCCGTCCTGCTG 300

      51 pLysSerThrGlnThrProSerProProCysGlnAlaPhe..... 64
          :::   ::::::   |||||||||   :::|||
     301 CCCGACCCCGCGAGCCCGGTCGCCACCGACGCGCGCTTCCAGGCGGCCCC 350

      65 .....................AsnHisTyrLeuSerAlaMetAlaSerIle 74
                              |||         :::|||   |||   :::
     351 CTGGCACTGGCGGGCCTGGGAACAGCTCACGCACAGCGTCCGCACCGGTG 400

      75 ArgGlnSerGlnGluGluProGluAspLeuArgProGluIleArgIleAl 91
         |||                           ||||||   :::|||   ::
     401 AGG.........................CGTCCTTTCCTTCGACGTGG 423

      91 aGln 92
         :|||
     424 CCAA 427

seq_name: /SIDS2/gcgdata/geneseq/geneseqn/NA2000.DAT:AAC55842

seq_documentation_block:
ID   AAC55842 standard; DNA; 53500 BP.
XX
AC   AAC55842;
XX
DT   19-JAN-2001  (first entry)
XX
DE   Complete nucleotide sequence of the mitomycin biosynthetic genes.
XX
KW   Mitomycin; biosynthesis; mitosane ring system; antibiotic; anti-cancer;
KW   anti-inflammatory; immune-enhancer; immunosuppressant; asthma;
KW   chronic obstructive pulmonary disease; respiratory inflammation;
KW   fungicide; pesticide; ds.
XX
OS   Streptomyces lavendulae.
XX
PN   WO200053737-A2.
XX
PD   14-SEP-2000.
XX
PF   10-MAR-2000; 2000WO-US06394.
XX
PR   12-MAR-1999;   99US-0266965.
XX
PA   (MINU ) UNIV MINNESOTA.
PA   (SHER/) SHERMAN D H.
PA   (MAOY/) MAO Y.
```

```
PA   (VARO/) VAROGLU M.
PA   (HEMM/) HE M.
PA   (SHEL/) SHELDON P C.
XX
PI   Sherman DH,  Mao Y,  Varoglu M,  He M,  Sheldon PC;
XX
DR   WPI; 2000-601980/57.
XX
PT   Novel nucleic acid molecule comprising mitomycin biosynthetic gene
PT   cluster useful for cloning mitomycin biosynthetic genes for elucidating
PT   the molecular basis of mitosane ring system biosynthesis -
XX
PS   Disclosure; Figure 26; 399pp; English.
XX
CC   This invention relates to isolated and purified nucleic acid molecules
CC   from the mitomycin biosynthetic gene cluster. Mitomycins are a group of
CC   natural products that contain a variety of functional groups, including
CC   amino benzoquinone and axiridine ring systems. The S. lavendulae
CC   mitomycin biosynthetic gene cluster comprises 47 mitomycin genes
CC   spanning 55kb of DNA. The invention includes an expression cassette
CC   comprising a mitomycin biosynthetic gene operably linked to a promoter,
CC   and host cells transformed with the cassette. The nucleotide, and protein
CC   sequences and the transformed host cells of the invention result in
CC   antiasthmatic, antiinflammatory, cytostatic, immunomodulatory, and
CC   antibiotic activities. The nucleotide sequences are used to elucidate the
CC   molecular basis for the biosynthesis of the mitosane ring system, as well
CC   as to engineer the biosynthesis of novel natural products, e.g.
CC   antibiotics, anti-inflammatory agents, anti-cancer agents,
CC   immune-enhancers, immunosuppressants, agents to treat asthma, chronic
CC   obstructive pulmonary disease as well as other disease involving
CC   respiratory inflammation, or cholesterol-lowering agents or as crop
CC   protection agents (e.g. fungicides or insecticides) as well as
CC   bipolymers, e.g., in packaging or biomedical applications, or to engineer
CC   PHA monomer synthases. Sequences AAC55782-C55881, AAC55815-C55849 and
CC   AAB32485-B32542 represent mitomycin biosynthetic gene cluster DNA
CC   sequences and encoded proteins. Sequences AAC55812-C55814,
CC   AAC55850-C55856 and AAC55862-C55869 represent PCR primers used in the
CC   cloning of the mitomycin biosynthetic genes.
XX
SQ   Sequence 53500 BP; 7481 A; 19740 C; 19126 G; 7153 T; 0 other;


alignment_scores:
           Quality:   92.50            Length:    118
             Ratio:  1.779              Gaps:      5
 Percent Similarity:  44.068   Percent Identity: 28.814

alignment_block:
 US-09-508-832-4 x AAC55842   ..

 Align seg 1/1  to: AAC55842  from: 1  to: 53500

       5 ProSerAsp...ValSerSerGluCysAspArgGluGlyGlyGlnLeuGl 20
         ||||||:::   :::|||   :::|||:::|||:::::
   42946 CCGTCGAACTCGGCGTCCCGGAACTGCTCCAGGAGGGCC..... 42984

      20 nProAlaGluArgProProGlnLeuArgProGlyAlaPro..... 33
          ||||||   |||||||||:::   ||||||   ||||||
   42985 .CCCGCACCGCGACCGCCCTCGCCGAGGCCACCGGCGCCCACGAGCAGAC 43033

      34 .....Thr 34
                                                 |||
   43034 CCTGCGCAGACTGCTCCGACTGCTCGCCACGGTGGGCGTCTTCGACGACC 43083

      35 SerLeuGlnThrGluProGlnAspArgSerProAlaProMetSerCysAs 51
         |||      |||      :::   |||:::|||:::|||   ||||||
   43084 TCGGCCACGACGACCTGTTCGCCCAGAACGCCCTCTCCGCCGTCCTGCTG 43133

      51 pLysSerThrGlnThrProSerProProCysGlnAlaPhe..... 64
           :::   ::::::   |||||||||   :::|||
   43134 CCCGACCCCGCGAGCCCGGTCGCCACCGACGCGCGCTTCCAGGCGGCCCC 43183

      65 .....AsnHisTyrLeuSerAlaMetAlaSerIle 74
                                  |||         :::|||   |||   :::
   43184 CTGGCACTGGCGGGCCTGGGAACAGCTCACGCACAGCGTCCGCACCGGTG 43233

      75 ArgGlnSerGlnGluGluProGluAspLeuArgProGluIleArgIleAl 91
         |||                        ||||||   :::|||   ::
   43234 AGG.....CGTCCTTTCCTTCGACGTGG 43256

      91 aGln 92
         :|||
   43257 CCAA 43260

seq_name: /SIDS2/gcgdata/geneseq/geneseqn/NA2001.DAT:AAS23095

seq_documentation_block:
ID   AAS23095 standard; DNA; 4452 BP.
XX
AC   AAS23095;
XX
DT   24-OCT-2001  (first entry)
XX
DE   DNA encoding novel bone marrow polypeptide #189.
XX
KW   Bone marrow; diagnostic; therapeutic; gene therapy; antigenic;
KW   haematopoiesis; myeloid; lymph cell disorder; tissue regeneration;
KW   wound healing; nutritional supplement; immune disorder;
KW   severe combined immunodeficiency; SCID; ds.
XX
OS   Homo sapiens.
XX
PN   WO200157187-A2.
XX
PD   09-AUG-2001.
XX
PF   05-FEB-2001; 2001WO-US03782.
XX
PR   03-FEB-2000; 2000US-0496914.
PR   20-JUN-2000; 2000US-0598075.
PR   19-JUL-2000; 2000US-0620325.
PR   30-NOV-2000; 2000US-0250683.
XX
PA   (HYSE-) HYSEQ INC.
XX
PI   Ford JE,  Boyle BJ,  Tang YT,  Liu C,  Asundi V,  Zhou P,  Xue AJ;
PI   Ren F,  Drmanac RT;
XX
DR   WPI; 2001-488875/53.
DR   P-PSDB; AAU14790.
XX
PT   Nucleic acids encoding bone marrow polypeptides, useful in diagnostic
PT   and gene therapy -
XX
PS   Claim 1; Page 358-361; 392pp; English.
XX
CC   AAS22907-AAS23099 represent nucleic acids encoding novel bone marrow
CC   polypeptides. The nucleic acids and corresponding proteins may be used
CC   in the prevention, diagnosis and treatment of diseases associated with
CC   inappropriate bone marrow polypeptide expression. For example, to treat
CC   disorders associated with decreased expression by rectifying mutations
CC   or deletions in a patient's genome that affect the activity of the
CC   polypeptides by expressing inactive proteins or to supplement the
CC   patient's own production of the polypeptide. Additionally, the nucleic
CC   acids may be used to produce the polypeptides, by inserting the nucleic
CC   acids into a host cell and culturing the cell to express the protein.
CC   The nucleic acid and its complementary sequences may also be used as DNA
CC   probes in diagnostic assays to detect and quantitate the presence of
CC   similar nucleic acid sequences in samples, and therefore which patients
CC   may be in need of restorative therapy. The proteins may also be used as
CC   antigens in the production of antibodies against bone marrow proteins
CC   and in assays to identify modulators of their expression and activity.
CC   The anti-bone marrow protein antibodies and antagonists may also be used
CC   to down regulate expression and activity. The antibodies may also be used
CC   as diagnostic agents for detecting the presence of the protein in samples
CC   (e.g. by enzyme linked immunosorbant assay (ELISA)). The proteins
```

```
CC   may be used to regulate haematopoiesis activity, and consequently in the
CC   treatment of myeloid or lymph cell disorders; in tissue regeneration,
CC   such as wound healing; as a nutritional supplement; and in treatment of
CC   immune disorders such as severe combined immunodeficiency (SCID).
XX
SQ   Sequence 4452 BP; 844 A; 1289 C; 1315 G; 1004 T; 0 other;


alignment_scores:
           Quality:   91.50            Length:    133
             Ratio:  1.386              Gaps:      6
 Percent Similarity:  49.624   Percent Identity: 27.068

alignment_block:
 US-09-508-832-4 x AAS23095   ..

 Align seg 1/1  to: AAS23095  from: 1  to: 4452

      17 GlyGlnLeuGlnProAlaGluArgProProGlnLeuArgProGlyAlaPr 33
         |||      ||||||   :::|||   ||||||   ||||||
     475 GGACCCCGCCAGCCCCAGCATCGCGCGCCGCAGCCGCGGCCC........ 516

      33 oThrSerLeuGlnThrGluProGlnAspArgSerProAlaProMetSerC 50
                                 ||||||   ||||||      :
     517 ..........................CGCAGCTCCGCCCCCGGCCCGG 538

      50 ysAspLysSerThrGlnThrProSerProProCysGlnAlaPheAsnHis 66
         ::        :::::::::|||:::   |||      |||   :::|||
     539 CCCGGCCCCGGGCCCGCTCGCCCGCCGCCCCGCATGGAGCTGTCAGCCAT 588

      67 TyrLeuSerAlaMetAlaSerIleArgGlnSerGlnGluGluProGluAs 83
                    |||:::::::|||::::::   |||:::||||||::
     589 .........CGGCGAGCAGGTGTTCGCCGTGGA...GAGCATCCGGAAGA 626

      83 pLeuArgProGlu......IleArgIleAlaGlnGluLeuArgArgIleG 98
         :   |||   |||         |||::::::   |||:::::::|||::::
     627 AGCGCGTGCGGAAGGGTAAAGTCGAGTATCTGGTGAAGTGGAAAGGATGG 676

      98 lyAspGluPheAsnGluThrTyrThrArgArgVal.Phe..........  110
         ::   :::   :::::::::   |||||||||::: :::
     677 CCCCCAAAGTACAGCACGTGGGAACCAGAAGAGCACATCTTGGACCCCCG 726

     111 .....................................AlaAsnAspT 114
                                              |||:::   |
     727 CCTCGTCATGGCCTACGAGGAGAAGGAGGAGAGAGACCGAGCATCGGGGT 776

     114 yrArgGluAlaGluAspHisProGlnMetValIleLeuGlnLeuLeu 129
         ||||||:::            |||:::   :::::::||||||   |||
     777 ATAGGAAGAGAGGTCCGAAACCCAAGCGGCTTCTGCTGCAGCGGCTG 823

seq_name: /SIDS2/gcgdata/geneseq/geneseqn/NA2001.DAT:AAF90327

seq_documentation_block:
ID   AAF90327 standard; DNA; 4033 BP.
XX
AC   AAF90327;
XX
DT   23-JUL-2001  (first entry)
XX
DE   Human JAFFA genomic DNA.
XX
KW   JAFFA; human; fibroblast growth factor; diagnosis; therapy;
KW   cancer; autoimmune disease; cytostatic; immunosuppressive;
KW   neuroprotective; ds.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   exon            1..168
FT                   /*tag=  a
FT                   /note= "exons 1 and 2 of the
FT                         (1,2)-fucosyltransferase gene"
FT   CDS             847..2794
FT                   /*tag=  b
FT                   /product= "JAFFA"
FT                   /note= "contains introns"
FT   exon            706..1081
FT                   /*tag=  c
FT                   /number= "1"
FT   intron          1081..1535
FT                   /*tag=  d
FT                   /number= "1"
FT   exon            1536..1639
FT                   /*tag=  e
FT                   /number= "2"
FT   intron          1640..2504
FT                   /*tag=  f
FT                   /number= "2"
FT   exon            2505..2792
FT                   /*tag=  g
FT                   /number= "3"
XX
PN   WO200138357-A2.
XX
PD   31-MAY-2001.
XX
PF   22-NOV-2000; 2000WO-US32181.
XX
PR   22-NOV-1999;   99US-0444165.
XX
PA   (MILL-) MILLENNIUM PHARM INC.
XX
PI   Khodadoust MM;
XX
DR   WPI; 2001-355881/37.
DR   P-PSDB; AAB82351.
XX
PT   Isolated JAFFA nucleic acid molecules which encode novel fibroblast
PT   growth factor family members (JAFFA) are useful for developing novel
PT   diagnostic and therapeutic agents for JAFFA-associated disorders such
PT   as cancers  -
XX
PS   Disclosure; Fig 6; 137pp; English.
XX
CC   The present sequence is that of the human JAFFA gene, a novel
CC   member of the fibroblast growth factor family.  Analysis of rare
CC   sequences in the Millennium database led to the identification of
CC   an expressed sequence tag from a human prostate tumour library which
CC   coded for a novel hypothetical signal peptide.  Further analysis
CC   led to the identification of JAFFA.  The invention provides JAFFA
CC   polypeptides, polynucleotides (see also AAF90326) and anti-JAFFA
CC   antibodies.  It also provides antisense nucleic acid molecules,
CC   recombinant expression vectors, host cells and transgenic animals.
CC   JAFFA proteins may be useful for developing novel diagnostic and
CC   therapeutic agents for JAFFA-associated disorders such as cancers
CC   and other disorders which demonstrate ABO(H) blood group disorders
CC   and for controlling cellular proliferative and/or differentiative
CC   disorders.  The JAFFA nucleic acid and protein may be used to treat
CC   and/or diagnose a variety of immune disorders such as autoimmune
CC   disease and multiple sclerosis.  JAFFA nucleic acids may be used to
CC   express JAFFA protein, to detect JAFFA mRNA or a genetic alteration
CC   in a JAFFA gene, and to modulate JAFFA activity.
XX
SQ   Sequence 4033 BP; 815 A; 1099 C; 1242 G; 877 T; 0 other;


alignment_scores:
           Quality:   90.50            Length:    109
             Ratio:  1.810              Gaps:      5
 Percent Similarity:  45.872   Percent Identity: 30.275

alignment_block:
 US-09-508-832-4 x AAF90327/rev   ..

 Align seg 1/1  to reverse of: AAF90327  from: 1  to: 4033
```

```
      5 ProSerAspValSerSerGluCysAspArgGluGlyGlyGlnLeuGlnPr 21
        |||||||||   ::::::          :::      :::||||||:::||
   1409 CCCTCAGACCCAGGAGAAGCCCAAGCCAAAGCCCCCTCCCAACTCAGACC 1360

     21 oAlaGluArgProProGln.......................LeuArgP 30
        |   :::      ||||||                       ||||
   1359 CAAGAGTCCAGGACCCCAGCCCTTCCTCCTTCAAACCCAAGAGTCCAGAC 1310

     30 roGlyAlaPro............ThrSerLeuGlnThrGluPro...... 40
        ||   ||||||              :::::||||||:::|||
   1309 CCCCAGCCCCTCCTCCCTCAGACCCAGGGGTCCAGACCCAGCCCCTCCTC 1260

     41 ............GlnAspArgSerProAlaProMetSerCysAspLysSe 53
                    ||||||||||||||||         |||   ::
   1259 CCTCAGACCCAGGGGCCCAGATCCCCAGCCCCTCCTCCCTCAGACCCAGG 1210

     53 rThrGlnThrProSerProProCysGlnAlaPheAsnHisTyrLeuSerA 70
        ::::|||   |||||||||
   1209 AGTCCAGATCCCCAGCCCC.............................. 1191

     70 laMetAlaSerIleArgGlnSerGlnGluGluProGluAspLeuArgPro 86
           :::|||:::|||         :::   |||:::   |||   |||
   1190 .....TCCTCCCTCAGACCCAGGGGCCAAGGCCCCCAGCCCCTCCTCCCT 1146

     87 GluIleArgIleAlaGlnGluLeuArg 95
        :::         :::||||||   |||
   1145 CAG.........ACCCAGGAATCCAGA 1128
```

seq_name: /SIDS2/gcgdata/geneseq/geneseqn/NA1999.DAT:AAZ32020

seq_documentation_block:
ID AAZ32020 standard; DNA; 38734 BP.
XX
AC AAZ32020;
XX
DT 10-JAN-2000 (first entry)
XX
DE Human METH1 related EST AL021529.
XX
KW Human; METH1; METH2; anti-angiogenic; metalloprotease thrombospondin;
KW cancer; diagnosis; hyperproliferative disorder; autoimmune disease;
KW angiogenesis inhibitor; abnormal wound healing; inflammation;
KW rheumatoid arthritis; psoriasis; endometrial bleeding disorder;
KW diabetic retinopathy; macula degeneration; haemangioma; detection;
KW arterial-venous malformation; immune deficiency; ss.
XX
OS Homo sapiens.
XX
PN WO9937660-A1.
XX
PD 29-JUL-1999.
XX
PF 22-JAN-1999; 99WO-US01313.
XX
PR 23-JAN-1998; 98US-0072298.
PR 28-AUG-1998; 98US-0098539.
XX
PA (IRUE/) IRUELA-ARISPE L.
PA (HAST/) HASTINGS G A.
PA (RUBE/) RUBEN S M.
XX
PI Iruela-Arispe L, Hastings GA, Ruben SM;
XX
DR WPI; 1999-590684/50.
XX
PT New isolated metalloprotease thrombospondin polypeptides, useful for
PT treating hyperproliferative disorders, cancers or autoimmune disorders
PT -
XX
PS Disclosure; Page 296-321; 457pp; English.
XX
CC AAZ32000 and AAZ32001 encode, and AAY49501 and AAY49502 represent, human
CC metalloprotease thrombospondin (METH) proteins METH1 and METH2
CC respectively. METH1 and METH2 have been found to be potent inhibitors of
CC angiogenesis both in vitro and in vivo. They can be used for treating
CC cancer and other disorders related to angiogenesis including abnormal
CC wound healing, inflammation, rheumatoid arthritis, psoriasis,
CC endometrial bleeding disorders, diabetic retinopathy, some forms of
CC macula degeneration, haemangiomas, and arterial-venous malformations.
CC They may be useful in treating deficiencies or disorders of the immune
CC system, by activating or inhibiting the proliferation, differentiation,
CC or mobilisation (chemotaxis) of immune cells. The etiology of these
CC immune deficiencies or disorders may be genetic, somatic, such as
CC cancer or some autoimmune disorders, acquired (e.g. by chemotherapy or
CC toxins), or infectious. They can also be used to treat inflammatory
CC conditions, both chronic and acute conditions. The products can also be
CC used for detection and diagnosis. AAZ32002 to AAZ32080, and AAY49503 to
CC AAY49511 represent sequences given in the exemplification of the present
CC invention.
XX
SQ Sequence 38734 BP; 6142 A; 13140 C; 13585 G; 5867 T; 0 other;

alignment_scores:
Quality: 90.50 Length: 136
Ratio: 1.437 Gaps: 6
Percent Similarity: 46.324 Percent Identity: 30.147

alignment_block:
US-09-508-832-4 x AAZ32020/rev ..

Align seg 1/1 to reverse of: AAZ32020 from: 1 to: 38734

```
      5 ProSerAspValSerSerGluCysAspArgGluGlyGlyGlnLeuGlnPr 21
        ||||||   :::::::|||   |||:::     :::   :::     ||
  26579 CCCTCGACCGCAACGTCGACCTGCTCGTACGCCGCTACCTCAGCGACCCC 26530

     21 oAlaGluArgProProGlnLeuArgProGly.................. 31
        |:::  .   |||         |||||||:::
  26529 GTCGTGCACTCCGTCGACCCCTCGGCCGGCGCCGTCACCACGATGGAGCA 26480

     32 ...................AlaProThrSerLeuGlnThrGluProGln 41
                            ||||||||||||      |||:::|||
  26479 CCACGTACTCCATGTCCACGGCGCCGACAAGC......ACGCGGCCACCA 26436

     42 AspArgSerPro.........AlaProMetSerCysAspLysSerThrGl 55
           |||||||||         |||      ||||||:::::::::   ::
  26435 CCGAGATCGCCGCACGCGACGGCCGCGTGATCATGTTCCTCGACACCAAG 26386

     55 nThrProSer....................................... 58
        :|||||||||
  26385 CACGCCGTCGACAGACTGACCCGGGACCTGCTCAACAGCGGGGTACGGGC 26336

     59 ..ProProCysGlnAlaPheAsnHisTyrLeuSerAlaMetAlaSerIle 74
          |||||||||   |||   :::         ||||||   |||
  26335 CGCCGCCCTGCACGGCGGCAAGTCCCAGCCGCAGCGCACCCGCACTCTGG 26286

     75 ArgGlnSerGln......GluGluProGluAspLeuArgProGluIleAr 89
        |||:::|||:::         :::|||        ||||||      ||
  26285 CGCAGTTCAAGACCGGGCACGTCACCGTGCTGGTGGCGACCAACGTCGCG 26236

     89 gIleAlaGlnGluLeuArgArgIleGlyAspGluPheAsnGluThrTyrT 106
        |   |||::::::             :::   :::   :::::::|||   |
  26235 GCACGCGGAATCCACGTCGACAACCTCGACCTCG...TCGTCAACGTCGA 26189

    106 hrArgArg 108
        ||||||||
  26188 CCCGCCGA 26181
```

seq_name: /SIDS2/gcgdata/geneseq/geneseqn/NA2001.DAT:AAC90077

seq_documentation_block:
ID AAC90077 standard; DNA; 38734 BP.
XX
AC AAC90077;
XX

```
DT   19-MAR-2001  (first entry)
XX
DE   AL021529 cDNA clone.
XX
KW   METH; metalloprotease; thrombospondin; angiogenesis inhibition;
KW   cancer therapy; benign tumour; ocular angiogenic disease;
KW   rheumatoid arthritis; psoriasis; wound healing; endometriosis;
KW   vasculogenesis; granulation; hypertrophic scar; nonunion fracture;
KW   scleroderma, trachoma; vascular adhesion; myocardial angiogenesis;
KW   coronary collateral; cerebral collateral; arteriovenous malformation;
KW   ischaemic limb angiogenesis; Osler-Webber syndrome; wound granulation;
KW   plaque neovascularisation; telangiectasia; haemophiliac joint; EST;
KW   angiofibroma; fibromuscular dysplasia; expressed sequence tag;
KW   Crohn's disease; atherosclerosis; birth control; ss.
XX
OS   Unidentified.
XX
PN   WO200071577-A1.
XX
PD   30-NOV-2000.
XX
PF   25-MAY-2000; 2000WO-US14462.
XX
PR   25-MAY-1999;   99US-0318208.
PR   20-JUL-1999;   99US-0144882.
PR   10-AUG-1999;   99US-0147823.
PR   13-AUG-1999;   99US-0373658.
PR   22-DEC-1999;   99US-0171503.
PR   22-FEB-2000; 2000US-0183792.
XX
PA   (HUMA-) HUMAN GENOME SCI INC.
PA   (SMIK ) SMITHKLINE BEECHAM CORP.
PA   (BETH-) BETH ISRAEL DEACONESS MEDICAL CENT.
PA   (IRUE/) IRUELA-ARISPE L.
PA   (HAST/) HASTINGS G A.
PA   (RUBE/) RUBEN S M.
PA   (JONA/) JONAK Z L.
PA   (TRUL/) TRULLI S H.
PA   (FORN/) FORNWALD J A.
PA   (TERR/) TERRETT J A.
XX
PI   Iruela-Arispe L,  Hastings GA,  Ruben SM,  Jonak ZL,  Trulli SH;
PI   Fornwald JA,  Terrett JA;
XX
DR   WPI; 2001-025136/03.
XX
PT   METH1 and METH2 polynucleotides and encoded polypeptides, used to
PT   inhibit angiogensis in the treatment of disorders such as cancer,
PT   rheumatoid arthritis and psoriasis -
XX
PS   Claim 7; Pages 597-622; 768pp; English.
XX
CC   The present invention relates to human METH1 and METH2, (ME for
CC   metalloprotease and TH for thrombospondin; see AAB50002 and AAB50003).
CC   The present sequence is an expressed sequence tag (EST) for METH. METH
CC   can be used for inhibiting angiogenesis in an individual, and for
CC   treating cancer, benign tumours, an ocular angiogenic disease,
CC   rheumatoid arthritis, psoriasis, delayed wound healing, endometriosis,
CC   vasculogenesis, granulations, hypertrophic scars, nonunion fractures,
CC   scleroderma, trachoma, vascular adhesions, myocardial angiogenesis,
CC   coronary collaterals, cerebral collaterals, arteriovenous malformations,
CC   ischaemic limb angiogenesis, Osler-Webber syndrome, plaque
CC   neovascularisation, telangiectasia, haemophiliac joints, angiofibroma,
CC   fibromuscular dysplasia, wound granulation, Crohn's disease or
CC   atherosclerosis. METH can also be used in birth control. METH can also
CC   be used in diagnostic methods for the prognosis of cancer.
XX
SQ   Sequence 38734 BP; 6142 A; 13138 C; 13586 G; 5868 T; 0 other;


alignment_scores:
           Quality:   90.50            Length:    136
             Ratio:  1.437              Gaps:      6

 Percent Similarity:  46.324   Percent Identity: 30.147

alignment_block:
 US-09-508-832-4 x AAC90077/rev   ..

 Align seg 1/1  to reverse of: AAC90077  from: 1  to: 38734

      5 ProSerAspValSerSerGluCysAspArgGluGlyGlyGlnLeuGlnPr 21
        ||||||  ::::::|||   |||:::      :::   :::      ||
   26579 CCCTCGACCGCAACGTCGACCTGCTCGTACGCCGCTACCTCAGCGACCCC 26530

     21 oAlaGluArgProProGlnLeuArgProGly..................... 31
        |:::      |||          ||||||:::
   26529 GTCGTGCACTCCGTCGACCCCTCGGCCGGCGCCGTCACCACGATGGAGCA 26480

     32 ...................AlaProThrSerLeuGlnThrGluProGln 41
                            ||||||||||||      |||:::|||
   26479 CCACGTACTCCATGTCCACGGCGCCGACAAGC......ACGCGGCCACCA 26436

     42 AspArgSerPro.........AlaProMetSerCysAspLysSerThrGl 55
          |||||||||          |||      ||||||:::::::::   ::
   26435 CCGAGATCGCCGCACGCGACGGCCGCGTGATCATGTTCCTCGACACCAAG 26386

     55 nThrProSer.......................................... 58
        :|||||||||
   26385 CACGCCGTCGACAGACTGACCCGGGACCTGCTCAACAGCGGGGTACGGGC 26336

     59 ..ProProCysGlnAlaPheAsnHisTyrLeuSerAlaMetAlaSerIle 74
          |||||||||   |||   :::          ||||||   |||
   26335 CGCCGCCCTGCACGGCGGCAAGTCCCAGCCGCAGCGCACCCGCACTCTGG 26286

     75 ArgGlnSerGln......GluGluProGluAspLeuArgProGluIleAr 89
        |||:::|||:::         :::|||         ||||||      ||
   26285 CGCAGTTCAAGACCGGGCACGTCACCGTGCTGGTGGCGACCAACGTCGCG 26236

     89 gIleAlaGlnGluLeuArgArgIleGlyAspGluPheAsnGluThrTyrT 106
        |   |||::::::           :::   :::   ::::::|||   |
   26235 GCACGCGGAATCCACGTCGACAACCTCGACCTCG...TCGTCAACGTCGA 26189

    106 hrArgArg 108
        ||||||||
   26188 CCCGCCGA 26181

seq_name: /SIDS2/gcgdata/geneseq/geneseqn/NA2001.DAT:AAI21994

seq_documentation_block:
ID   AAI21994 standard; DNA; 3331 BP.
XX
AC   AAI21994;
XX
DT   12-OCT-2001  (first entry)
XX
DE   Probe #11927 for gene expression analysis in human cervical cell sample.
XX
KW   Probe; human; microarray; gene expression; cervical epithelial cell;
KW   cervical cancer; ss.
XX
OS   Homo sapiens.
XX
PN   WO200157278-A2.
XX
PD   09-AUG-2001.
XX
PF   30-JAN-2001; 2001WO-US00670.
XX
PR   04-FEB-2000; 2000US-0180312.
PR   26-MAY-2000; 2000US-0207456.
PR   30-JUN-2000; 2000US-0608408.
PR   03-AUG-2000; 2000US-0632366.
PR   21-SEP-2000; 2000US-0234687.
PR   27-SEP-2000; 2000US-0236359.
PR   04-OCT-2000; 2000GB-0024263.
XX
```

```
PA   (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI   Penn SG,  Hanzel DK,  Chen W,  Rank DR;
XX
DR   WPI; 2001-488901/53.
XX
PT   Human genome-derived single exon nucleic acid probes useful for
PT   analyzing gene expression in human cervical epithelial cells -
XX
PS   Claim 25; SEQ ID No 11927; 487pp; English.
XX
CC   The present invention relates to human single exon nucleic acid probes
CC   (SENP). The present sequence is one such probe. The SENPs are derived
CC   from human HeLa cells. The SENPs can be used to produce a single exon
CC   microarray, which can be used for measuring human gene expression in a
CC   sample derived from human cervical epithelial cells. By measuring gene
CC   expression, the probes are therefore useful in grading and/or staging
CC   of diseases of the cervix, notably cervical cancer.
CC   Note: The sequence data for this patent did not form part of the printed
CC   specification, but was obtained in electronic format directly from WIPO
CC   at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ   Sequence 3331 BP; 660 A; 619 C; 1183 G; 869 T; 0 other;


alignment_scores:
            Quality:   89.00              Length:    134
              Ratio:  1.328                Gaps:      3
 Percent Similarity:  50.000   Percent Identity: 27.612

alignment_block:
 US-09-508-832-4 x AAI21994/rev   ..

 Align seg 1/1  to reverse of: AAI21994  from: 1  to: 3331

       2 AlaLysGlnProSerAspValSerSerGluCysAspArgGluGlyGlyGl 18
         |||   :::|||:::::::::      ::::::   |||
    2967 GCTCTGAGACCAGCACAGATTCTACCACAGGCTCTGAGACCACCACAGCC 2918

      18 nLeuGlnProAlaGluArgProProGlnLeuArgProGlyAlaProThrS 35
          |||      |||   ||||||||||||||||      |||
    2917 TCTACTGAAGGCTCTGAGACCACCACAGCTTCCACTGAAGGCTCTGAGGC 2868

      35 erLeuGlnThrGluProGlnAspArgSerProAlaProMetSerCysAsp 51
           ||||||:::   ||||||      :::      |||::::::
    2867 CACTACAGTCTCCACCACAG......GCTCTGAGACCACTACAGTTTCTA 2824

      52 LysSerThrGlnThrProSerProProCysGlnAlaPheAsnHisTyrLe 68
         ::::::||||||   |||   ||||||
    2823 TCACAGACTCAGAGACCACCACCACCT..................... 2797

      68 uSerAlaMetAlaSerIleArgGlnSerGlnGluGluProGluAspLeuA 85
             ::::::   ::::::|||      |||:::   |||::::::||||
    2796 ....GTACTGAAGGCTCTGAGATGACTGCAGTCTCCACCACAGTCTTTGA 2751

      85 rgProGluIleArgIleAlaGlnGluLeuArgArgIleGlyAspGluPhe 101
         |||||         :::   :::   ||||||              :::
    2750 GACCACTACAGCCTCTACTGAAGGCTCTGAGATCACAATAGCCTCTACTT 2701

     102 AsnGluThrTyrThrArgArgValPheAlaAsnAspTyrArgGluAlaGl 118
         :::               :::::::::   :::      |||      ::
    2700 CAGACTCTGAGACCACCACAGCTTCTACTGAAGGTTCTGAGACCACTACA 2651

     118 uAspHisProGlnMetVal.............IleLeuGlnLeuLeuA 130
         ::::   ||||||   :::             |||||||||   ||||
    2650 GTCACTACCGCAGGCTCTGAGACCAAAACAGCCTATACTACAGGCTCTGA 2601

     130 rg 130
         ||
    2600 GA 2599

seq_name: /SIDS2/gcgdata/geneseq/geneseqn/NA2001.DAT:AAI47284

seq_documentation_block:
ID   AAI47284 standard; DNA; 3331 BP.
XX
AC   AAI47284;
XX
DT   17-OCT-2001  (first entry)
XX
DE   Probe #15970 used to measure gene expression in human placenta sample.
XX
KW   Probe; microarray; human; placenta; antenatal diagnosis;
KW   genetic disorder; ss.
XX
OS   Homo sapiens.
XX
PN   WO200157272-A2.
XX
PD   09-AUG-2001.
XX
PF   30-JAN-2001; 2001WO-US00663.
XX
PR   04-FEB-2000; 2000US-0180312.
PR   26-MAY-2000; 2000US-0207456.
PR   30-JUN-2000; 2000US-0608408.
PR   03-AUG-2000; 2000US-0632366.
PR   21-SEP-2000; 2000US-0234687.
PR   27-SEP-2000; 2000US-0236359.
PR   04-OCT-2000; 2000GB-0024263.
XX
PA   (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI   Penn SG,  Hanzel DK,  Chen W,  Rank DR;
XX
DR   WPI; 2001-488897/53.
XX
PT   Human genome-derived single exon nucleic acid probes useful for
PT   analyzing gene expression in human placenta -
XX
PS   Claim 25; SEQ ID No 15970; 654pp; English.
XX
CC   The present invention relates to single exon nucleic acid probes (SENP).
CC   The present sequence is one such probe. The probes are useful for
CC   producing a microarray for predicting, measuring and displaying gene
CC   expression in samples derived from human placenta. The probes are useful
CC   for antenatal diagnosis of human genetic disorders.
XX
SQ   Sequence 3331 BP; 660 A; 619 C; 1183 G; 869 T; 0 other;


alignment_scores:
            Quality:   89.00              Length:    134
              Ratio:  1.328                Gaps:      3
 Percent Similarity:  50.000   Percent Identity: 27.612

alignment_block:
 US-09-508-832-4 x AAI47284/rev   ..

 Align seg 1/1  to reverse of: AAI47284  from: 1  to: 3331

       2 AlaLysGlnProSerAspValSerSerGluCysAspArgGluGlyGlyGl 18
         |||   :::|||:::::::::      ::::::   |||
    2967 GCTCTGAGACCAGCACAGATTCTACCACAGGCTCTGAGACCACCACAGCC 2918

      18 nLeuGlnProAlaGluArgProProGlnLeuArgProGlyAlaProThrS 35
          |||      |||   ||||||||||||||||      |||
    2917 TCTACTGAAGGCTCTGAGACCACCACAGCTTCCACTGAAGGCTCTGAGGC 2868

      35 erLeuGlnThrGluProGlnAspArgSerProAlaProMetSerCysAsp 51
           ||||||:::   ||||||      :::      |||::::::
    2867 CACTACAGTCTCCACCACAG......GCTCTGAGACCACTACAGTTTCTA 2824

      52 LysSerThrGlnThrProSerProProCysGlnAlaPheAsnHisTyrLe 68
         ::::::||||||   |||   ||||||
    2823 TCACAGACTCAGAGACCACCACCACCT..................... 2797
```

```
     68 uSerAlaMetAlaSerIleArgGlnSerGlnGluGluProGluAspLeuA 85
        ::::::   :::::::|||      |||:::   |||::::::||||
   2796 ....GTACTGAAGGCTCTGAGATGACTGCAGTCTCCACCACAGTCTTTGA 2751

     85 rgProGluIleArgIleAlaGlnGluLeuArgArgIleGlyAspGluPhe 101
        |||||         :::   :::   ||||||              :::
   2750 GACCACTACAGCCTCTACTGAAGGCTCTGAGATCACAATAGCCTCTACTT 2701

    102 AsnGluThrTyrThrArgArgValPheAlaAsnAspTyrArgGluAlaGl 118
        :::              :::::::::   :::      |||      ::
   2700 CAGACTCTGAGACCACCACAGCTTCTACTGAAGGTTCTGAGACCACTACA 2651

    118 uAspHisProGlnMetVal.............IleLeuGlnLeuLeuA 130
        ::::   ||||||   :::              |||||||||   ||||
   2650 GTCACTACCGCAGGCTCTGAGACCAAAACAGCCTATACTACAGGCTCTGA 2601

    130 rg 130
        ||
   2600 GA 2599

seq_name: /SIDS2/gcgdata/geneseq/geneseqn/NA2001.DAT:AAI07689

seq_documentation_block:
ID   AAI07689 standard; DNA; 3331 BP.
XX
AC   AAI07689;
XX
DT   09-OCT-2001  (first entry)
XX
DE   Probe #7680 used to measure gene expression in human breast sample.
XX
KW   Probe; human; breast disease; breast cancer; development disorder; ss;
KW   inflammatory disease; proliferative breast disease; non-carcinoma tumour.
XX
OS   Homo sapiens.
XX
PN   WO200157270-A2.
XX
PD   09-AUG-2001.
XX
PF   29-JAN-2001; 2001WO-US00661.
XX
PR   04-FEB-2000; 2000US-0180312.
PR   26-MAY-2000; 2000US-0207456.
PR   30-JUN-2000; 2000US-0608408.
PR   03-AUG-2000; 2000US-0632366.
PR   21-SEP-2000; 2000US-0234687.
PR   27-SEP-2000; 2000US-0236359.
PR   04-OCT-2000; 2000GB-0024263.
XX
PA   (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI   Penn SG,  Hanzel DK,  Chen W,  Rank DR;
XX
DR   WPI; 2001-476286/51.
XX
PT   Novel single exon nucleic acid probe used to measuring gene expression
PT   in a human breast -
XX
PS   Claim 25; SEQ ID No 7680; 322pp; English.
XX
CC   The present invention relates to novel single exon nucleic acid probes.
CC   The present sequence is one such probe. The probes are useful for
CC   measuring human gene expression in a human breast sample, where the probe
CC   hybridises at high stringency to a nucleic acid expressed in the human
CC   breast. The probes are useful for predicting, diagnosing, grading,
CC   staging, monitoring and prognosing diseases of the human breast,
CC   particularly those diseases with polygenic aetiology. The diseases
CC   include: breast cancer, disorders of development, inflammatory diseases
CC   of the breast, fibrocystic changes, proliferative breast disease and
CC   non-carcinoma tumours.
CC   Note: The sequence data for this patent did not form part of the printed
CC   specification, but was obtained in electronic format directly from WIPO
CC   at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ   Sequence 3331 BP; 660 A; 619 C; 1183 G; 869 T; 0 other;

alignment_scores:
           Quality:   89.00              Length:    134
             Ratio:  1.328                Gaps:      3
 Percent Similarity:  50.000   Percent Identity: 27.612

alignment_block:
 US-09-508-832-4 x AAI07689/rev   ..

 Align seg 1/1  to reverse of: AAI07689  from: 1  to: 3331

      2 AlaLysGlnProSerAspValSerSerGluCysAspArgGluGlyGlyGl 18
        |||   :::|||:::::::::      ::::::   |||
   2967 GCTCTGAGACCAGCACAGATTCTACCACAGGCTCTGAGACCACCACAGCC 2918

     18 nLeuGlnProAlaGluArgProProGlnLeuArgProGlyAlaProThrS 35
        |||      |||   ||||||||||||||||         |||
   2917 TCTACTGAAGGCTCTGAGACCACCACAGCTTCCACTGAAGGCTCTGAGGC 2868

     35 erLeuGlnThrGluProGlnAspArgSerProAlaProMetSerCysAsp 51
        ||||||:::   ||||||      :::      |||::::::
   2867 CACTACAGTCTCCACCACAG......GCTCTGAGACCACTACAGTTTCTA 2824

     52 LysSerThrGlnThrProSerProProCysGlnAlaPheAsnHisTyrLe 68
        ::::::||||||   |||   ||||||
   2823 TCACAGACTCAGAGACCACCACCACCT..................... 2797

     68 uSerAlaMetAlaSerIleArgGlnSerGlnGluGluProGluAspLeuA 85
        ::::::   ::::::|||      |||:::   |||::::::||||
   2796 ....GTACTGAAGGCTCTGAGATGACTGCAGTCTCCACCACAGTCTTTGA 2751

     85 rgProGluIleArgIleAlaGlnGluLeuArgArgIleGlyAspGluPhe 101
        |||||         :::   :::   ||||||              :::
   2750 GACCACTACAGCCTCTACTGAAGGCTCTGAGATCACAATAGCCTCTACTT 2701

    102 AsnGluThrTyrThrArgArgValPheAlaAsnAspTyrArgGluAlaGl 118
        :::              :::::::::   :::      |||      ::
   2700 CAGACTCTGAGACCACCACAGCTTCTACTGAAGGTTCTGAGACCACTACA 2651

    118 uAspHisProGlnMetVal.............IleLeuGlnLeuLeuA 130
        ::::   ||||||   :::              |||||||||   ||||
   2650 GTCACTACCGCAGGCTCTGAGACCAAAACAGCCTATACTACAGGCTCTGA 2601

    130 rg 130
        ||
   2600 GA 2599

seq_name: /SIDS2/gcgdata/geneseq/geneseqn/NA1999.DAT:AAZ32025

seq_documentation_block:
ID   AAZ32025 standard; DNA; 9810 BP.
XX
AC   AAZ32025;
XX
DT   10-JAN-2000  (first entry)
XX
DE   Human METH1 related EST AF018073.
XX
KW   Human; METH1; METH2; anti-angiogenic; metalloprotease thrombospondin;
KW   cancer; diagnosis; hyperproliferative disorder; autoimmune disease;
KW   angiogenesis inhibitor; abnormal wound healing; inflammation;
KW   rheumatoid arthritis; psoriasis; endometrial bleeding disorder;
KW   diabetic retinopathy; macula degeneration; haemangioma; detection;
KW   arterial-venous malformation; immune deficiency; ss.
XX
OS   Homo sapiens.
XX
PN   WO9937660-A1.
```

```
XX
PD   29-JUL-1999.
XX
PF   22-JAN-1999;   99WO-US01313.
XX
PR   23-JAN-1998;   98US-0072298.
PR   28-AUG-1998;   98US-0098539.
XX
PA   (IRUE/) IRUELA-ARISPE L.
PA   (HAST/) HASTINGS G A.
PA   (RUBE/) RUBEN S M.
XX
PI   Iruela-Arispe L,  Hastings GA,  Ruben SM;
XX
DR   WPI; 1999-590684/50.
XX
PT   New isolated metalloprotease thrombospondin polypeptides, useful for
PT   treating hyperproliferative disorders, cancers or autoimmune disorders
PT   -
XX
PS   Disclosure; Page 353-359; 457pp; English.
XX
CC   AAZ32000 and AAZ32001 encode, and AAY49501 and AAY49502 represent, human
CC   metalloprotease thrombospondin (METH) proteins METH1 and METH2
CC   respectively. METH1 and METH2 have been found to be potent inhibitors of
CC   angiogenesis both in vitro and in vivo. They can be used for treating
CC   cancer and other disorders related to angiogenesis including abnormal
CC   wound healing, inflammation, rheumatoid arthritis, psoriasis,
CC   endometrial bleeding disorders, diabetic retinopathy, some forms of
CC   macula degeneration, haemangiomas, and arterial-venous malformations.
CC   They may be useful in treating deficiencies or disorders of the immune
CC   system, by activating or inhibiting the proliferation, differentiation,
CC   gr mobilisation (chemotaxis) of immune cells. The etiology of these
CC   immune deficiencies or disorders may be genetic, somatic, such as
CC   cancer or some autoimmune disorders, acquired (e.g. by chemotherapy or
CC   toxins), or infectious. They can also be used to treat inflammatory
CC   conditions, both chronic and acute conditions. The products can also be
CC   used for detection and diagnosis. AAZ32002 to AAZ32080, and AAY49503 to
CC   AAY49511 represent sequences given in the exemplification of the present
CC   invention.
XX
SQ   Sequence 9810 BP; 1583 A; 3401 C; 3201 G; 1625 T; 0 other;


alignment_scores:
            Quality:   89.00              Length:    114
              Ratio:  1.679                Gaps:      5
 Percent Similarity:  46.491   Percent Identity: 28.947

alignment_block:
 US-09-508-832-4 x AAZ32025/rev   ..

 Align seg 1/1  to reverse of: AAZ32025  from: 1  to: 9810

       4 GlnProSerAspValSerSerGluCysAspArgGluGlyGlyGlnLeuGl 20
         :::|||   :::      |||      ||||||:::      :::   ::
    8050 CGGCCGCCAGAAGATCGTTCTGCGCCAGATCGCCGTAGACTTCGGTCATC 8001

      20 nProAlaGluArgProProGlnLeuArgProGlyAlaProThrSerLeuG 37
         :||||||   ||||||   :::      |||:::||||||
    8000 GCCAGCCACTCGGCCGGCGTCTCCTTCGCCCGCCGCGCCCGGTCCTGCAG 7951

      37 lnThrGluProGlnAspArgSerProAlaProMetSerCysAspLysSer 53
            :::   :::|||||||||      |||
    7950 CGCCGTCCAGTTCGGATCGTTCGGCTCGACCA..................... 7919

      54 ThrGlnThrProSerProProCys.GlnAlaPheAsnHisTyrLeuSerA 70
         |||   |||||||||   ||| :::|||                    :
    7918 ...CAACGCCCGAGTCCGTCGTGCCGAAGCA.................G 7890

      70 laMetAlaSerIleArgGlnSerGlnGluGluPro.............. 81
         ::   |||      |||:::::::|||:::::::|||
    7889 TAGCGGCACCAGAGCGCCGAGAGCAGCACCAGCCCCTTCGGCACCGTGCC 7840

      82 ..........................GluAspLeuArgProGluIl 88
                                     ::::::||||||      ::
    7839 CGCCGCCAGATTGTCGCGCAGCGACGGCACGATGAACTTCGGCTGCCGGT 7790

      88 eArgIle......AlaGlnGluLeuArgArgIleGlyAsp 99
         :|||         |||:::      ||||||:::||||||
    7789 TCGAACCGTCGAGGCAGAGCCTGCGCGTCGTGTCGGCGAT 7750
```

OM of: US-09-508-832-4 to: EST:* out_format : pfs

Date: Dec 11, 2001 1:03 AM

About: Results were produced by the GenCore software, version 4.5,
Copyright (c) 1993-2000 Compugen Ltd.

Command line parameters:
-MODEL=frame+_p2n.model -DEV=xlp
-Q=/cgn2_1/USPTO_spool/US09508832/runat_10122001_110349_29536/app_query.fasta_1.620
-DB=EST -QFMT=fastap -SUFFIX=rst -GAPOP=12.000 -GAPEXT=4.000
-MINMATCH=0.100 -LOOPCL=0.000 -LOOPEXT=0.000 -QGAPOP=4.500
-QGAPEXT=0.050 -XGAPOP=10.000 -XGAPEXT=0.500 -FGAPOP=6.000
-FGAPEXT=7.000 -YGAPOP=10.000 -YGAPEXT=0.500 -DELOP=6.000
-DELEXT=7.000 -START=1 -MATRIX=blosum62 -TRANS=human40.cdi
-LIST=45 -DOCALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0
-ALIGN=15 -MODE=LOCAL -OUTFMT=pfs -NORM=ext -MINLEN=0
-MAXLEN=2000000000 -USER=US09508832_@CGN1_1_0 -NCPU=6 -ICPU=3
-LONGLOG -NO_XLPXY -WAIT -THREADS=1

Search information block:
Query: US-09-508-832-4
Query length: 140
Database: EST:*
Database sequences: 11351937
Database length: 1077921985
Search time (sec): 2629.110000

score_list:

Sequence	Strd	Orig	ZScore	EScore	Len	! Documentation ..
gb_htc:AK011490	+	675.50	953.68	5.9e-44	1206	! AK011490 Mus musculus 10 days
gb_est2:BG921698	+	583.00	824.88	8.8e-37	935	! BG921698 602825518F1 NCI_CGAP_M
gb_est2:BF021882	+	525.00	748.90	1.5e-32	452	! BF021882 uy59b09.y1 McCarrey Ec
gb_est1:AI971169	-	370.00	528.77	2.7e-20	492	! AI971169 wr24h12.x1 NCI_CGAP_Pr
gb_est2:BG173095	+	354.00	503.54	7.0e-19	668	! BG173095 602336666F1 NCI_CGAP_M
gb_gss:AZ706148	+	347.00	494.82	2.1e-18	580	! AZ706148 RPCI-23-227P3.TV RPCI-
gb_est1:AW629314	-	287.00	408.75	1.3e-13	664	! AW629314 hi56e02.x1 Soares_NFL_
gb_est2:BF319454	-	249.00	359.46	7.4e-11	389	! BF319454 uy59b09.x1 McCarrey Ec
gb_est1:AA629308	-	127.00	184.63	0.4043	501	! AA629308 zu84g06.s1 Soares_test
gb_est2:BF172831	+	126.00	190.53	0.1895	210	! BF172831 PCL5805 Myeloma (PCL)
gb_est1:AA629050	-	123.00	178.96	0.8358	501	! AA629050 zu84a06.s1 Soares_test
gb_est1:AF209718	+	98.50	143.70	76.99	537	! AF209718 AF209718 Xenopus laevi
gb_est2:BG649655	+	98.00	145.84	58.52	383	! BG649655 RHIZ2_82_D01.b1_A003 R
gb_est2:BG412166	+	98.00	143.28	81.26	519	! BG412166 OV2_38_F11.b1_A002 Ova
gb_est1:BE728552	+	95.50	136.55	192.62	758	! BE728552 601562025F1 NIH_MGC_20
gb_est2:BE903772	+	95.50	135.66	215.77	842	! BE903772 601676601F1 NIH_MGC_21
gb_est1:BE563554	+	95.50	134.99	235.21	912	! BE563554 601334851F1 NIH_MGC_39
gb_est2:BE910153	+	95.50	132.29	332.35	1256	! BE910153 601503441F1 NIH_MGC_7
gb_est1:BE622003	+	95.50	132.19	336.92	1272	! BE622003 601440653F1 NIH_MGC_7
gb_est1:AL542837	-	95.00	133.99	267.34	944	! AL542837 AL542837 LTI_FL002_PL1
gb_est1:BE416943	+	94.50	137.06	180.40	603	! BE416943 MUG014.H03F990622 ITEC
gb_est1:BE416944	+	94.50	137.06	180.40	603	! BE416944 MUG014.H03R990622 ITEC
gb_est2:BG713458	+	94.50	136.80	186.54	622	! BG713458 pg11n.pk004.i5 Normali
gb_est1:BE394838	+	93.50	133.67	278.52	762	! BE394838 601312080F1 NIH_MGC_44
gb_est1:BE732875	+	93.50	132.56	321.39	870	! BE732875 601567809F1 NIH_MGC_21
gb_gss:AZ137807	+	93.50	132.25	334.18	902	! AZ137807 SP_0180_A2_H09_SP6E St
gb_est2:BF213947	-	93.50	132.02	344.20	927	! BF213947 601847413F1 NIH_MGC_55
gb_est2:BG913808	-	93.00	132.35	330.14	820	! BG913808 602812986F1 NCI_CGAP_E
gb_est2:BG337043	+	93.00	131.56	365.06	900	! BG337043 602434310F1 NIH_MGC_46
gb_est2:BG699551	+	92.50	134.34	255.66	595	! BG699551 602679256F1 NIH_MGC_9
gb_est2:BI394478	+	92.50	133.33	291.11	671	! BI394478 pgp1n.pk014.m20 Normal
gb_est1:BE394827	+	92.50	132.31	331.60	757	! BE394827 601312068F1 NIH_MGC_44
gb_est2:BI260770	+	92.50	132.21	335.87	766	! BI260770 602970728F1 NIH_MGC_12
gb_est2:BE799245	+	92.50	131.45	370.57	839	! BE799245 601586889F1 NIH_MGC_7
gb_gss:CNS044MO	-	92.50	130.51	418.01	938	! AL274281 Tetraodon nigroviridis
gb_est2:BG774733	+	92.50	127.35	626.87	1365	! BG774733 602662784F1 NIH_MGC_2
gb_est2:BG990772	-	92.00	139.71	128.33	289	! BG990772 MR1-HT1136-260101-015-
gb_est2:BF620177	-	92.00	133.69	277.92	591	! BF620177 HVSMEc0018M03f Hordeum
gb_gss:CNS01T4M	-	92.00	128.67	528.76	1072	! AL166063 Tetraodon nigroviridi
gb_gss:AZ414049	-	91.50	133.39	288.79	563	! AZ414049 1M0188013F Mouse 10kb
gb_est1:AL507998	+	91.50	131.92	348.50	670	! AL507998 AL507998 Hordeum vulga
gb_est2:BG752199	+	91.50	129.13	498.92	934	! BG752199 602731331F1 NIH_MGC
gb_est2:BF866948	+	91.50	124.51	902.01	1616	! BF866948 963082E01.y1 C. re
gb_est2:BF727326	+	91.00	132.22	335.69	595	! BF727326 by20a03.y1 Human Le
gb_est1:BE366432	-	91.00	131.77	355.23	627	! BE366432 PI1_33_A12.g1_A002

seq_name: gb_htc:AK011490

seq_documentation_block:

LOCUS AK011490 1206 bp mRNA HTC 05-JUL-2001
DEFINITION Mus musculus 10 days embryo cDNA, RIKEN full-length enriched library, clone:2610020M23, full insert sequence.
ACCESSION AK011490
VERSION AK011490.1 GI:12847647
KEYWORDS CAP trapper.
SOURCE Mus musculus (strain:C57BL/6J) 10 days embryo cDNA to mRNA, clone_lib:RIKEN full-length enriched mouse cDNA library clone:2610020M23.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 1206)
AUTHORS Carninci,P. and Hayashizaki,Y.
TITLE High-efficiency full-length cDNA cloning
JOURNAL Methods in enzymology. 303, 19-44 (1999)
MEDLINE 99279253
PUBMED 10349636
REFERENCE 2 (bases 1 to 1206)
AUTHORS Carninci,P., Shibata,Y., Hayatsu,N., Sugahara,Y., Shibata,K., Itoh,M., Konno,H., Okazaki,Y., Muramatsu,M. and Hayashizaki,Y.
TITLE Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes
JOURNAL Genome research. 10 (10), 1617-1630 (2000)
MEDLINE 20499374
PUBMED 11042159
REFERENCE 3 (bases 1 to 1206)
AUTHORS Shibata,K., Itoh,M., Aizawa,K., Nagaoka,S., Sasaki,N., Carninci,P., Konno,H., Akiyama,J., Nishi,K., Kitsunai,T., Tashiro,H., Itoh,M., Sumi,N., Ishii,Y., Nakamura,S., Hazama,M., Nishine,T., Harada,A., Yamamoto,R., Matsumoto,H., Sakaguchi,S., Ikegami,T., Kashiwagi,K., Fujiwake,S., Inoue,K., Togawa,Y., Izawa,M., Ohara,E., Watahiki,M., Yoneda,Y., Ishikawa,T., Ozawa,K., Tanaka,T., Matsuura,S., Kawai,J., Okazaki,Y., Muramatsu,M., Inoue,Y., Kira,A. and Hayashizaki,Y.
TITLE RIKEN integrated sequence analysis (RISA) system--384-format sequencing pipeline with 384 multicapillary sequencer
JOURNAL Genome research. 10 (11), 1757-1771 (2000)
MEDLINE 20530913
PUBMED 11076861
REFERENCE 4 (bases 1 to 1206)
AUTHORS The RIKEN Genome Exploration Research Group Phase II Team and the FANTOM Consortium.
TITLE Functional annotation of a full-length mouse cDNA collection
JOURNAL Nature 409, 685-690 (2001)
REFERENCE 5 (bases 1 to 1206)
AUTHORS Adachi,J., Aizawa,K., Akahira,S., Akimura,T., Aono,H., Arai,A., Arakawa,T., Carninci,P., Fukuda,S., Fukunishi,Y., Furuno,M., Hanagaki,T., Hara,A., Hayatsu,N., Hiramoto,K., Hiraoka,T., Hori,F., Imotani,K., Ishii,Y., Itoh,M., Izawa,M., Kato,H., Kawai,J., Kojima,Y., Konno,H., Kouda,M., Koya,S., Kurihara,C., Matsuyama,T., Miyazaki,A., Nishi,K., Nomura,K., Numazaki,R., Ohno,M., Okazaki,Y., Okido,T., Owa,C., Saito,H., Saito,R., Sakai,C., Sakai,K., Sano,H., Sasaki,D., Shibata,K., Shibata,Y., Shinagawa,A., Shiraki,T., Sogabe,Y., Suzuki,H., Tagami,M., Tagawa,A., Takahashi,F., Tanaka,T., Tejima,Y., Toya,T., Yamamura,T., Yasunishi,A., Yoshida,K., Yoshino,M., Muramatsu,M. and Hayashizaki,Y.
TITLE Direct Submission
JOURNAL Submitted (10-JUL-2000) Yoshihide Hayashizaki, The Institute of Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration Research Group, RIKEN Gemomic Sciences Center (GSC), RIKEN Yokohama Institute; 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan (E-mail:genome-res@gsc.riken.go.jp, URL:http://genome.gsc.riken.go.jp/, Tel:81-45-503-9222, Fax:81-45-503-9216)
COMMENT Please visit our web site (http://genome.gsc.riken.go.jp/) for

```
          further details.
          cDNA library was prepared and sequenced in Mouse Genome
          Encyclopedia Project of Genome Exploration Research Group in Riken
          Genomic Sciences Center and Genome Science Laboratory in RIKEN.
          Division of Experimental Animal Research in Riken contributed to
          prepare mouse tissues. First strand cDNA was primed with a primer
          [5' GAGAGAGAGAAGGATCCAAGAGCTCTTTTTTTTTTTTTTTVN 3'], cDNA was
          prepared by using trehalose thermo-activated reverse transcriptase
          and subsequently enriched for full-length by cap-trapper. cDNA went
          through one round of normalization to Rot = 7.5 and subtraction to
          Rot = 37.5. Second strand cDNA was prepared with the primer adapter
          of sequence [5'
          GAGAGAGAGATTCTCGAGTTAATTAAATTAATCCCCCCCCCCCCC 3']. cDNA was cleaved
          with XhoI and SstI. Cloning sites, 5' end: XhoI; 3' end: SstI.
          Host: SOLR.
FEATURES             Location/Qualifiers
     source          1. .1206
                     /organism="Mus musculus"
                     /strain="C57BL/6J"
                     /db_xref="taxon:10090"
                     /db_xref="MGD:MGI:1197519"
                     /db_xref="MGD:MGI:1902115"
                     /clone="2610020M23"
                     /clone_lib="RIKEN full-length enriched mouse cDNA library"
                     /dev_stage="10 days embryo"
BASE COUNT      265 a    339 c    298 g    301 t      3 others
ORIGIN


alignment_scores:
            Quality:  675.50            Length:    196
              Ratio:  4.895               Gaps:      2
 Percent Similarity:  70.408   Percent Identity: 69.898

alignment_block:
 US-09-508-832-4 x AK011490   ..

 Align seg 1/1  to: AK011490  from: 1  to: 1206

       1 MetAlaLysGlnProSerAspValSerSerGluCysAspArgGluGlyGl 17
         ||||||||||||||||||||||||   ||||||||||||   ||||||||
     220 ATGGCCAAGCAACCTTCTGATGTACGTTCTGAGTGTGACAGCGAAGGTGG 269

      17 yGlnLeuGlnProAlaGluArgProProGlnLeuArgProGlyAlaProT 34
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     270 ACAATTGCAGCCTGCTGAGAGGCCTCCCCAGCTCAGGCCTGGGGCCCCTA 319

      34 hrSerLeuGlnThrGluPro..............................  40
         ||||||||||||||||||
     320 CCTCCCTACAGACAGAACCGCAACGTAATCCGACGGGGAAGGGGACCGTG 369

      40 ..................................................  40

     370 CCCCCACGGCAGCCCTCAGGGCCCGCTGGCCCCACCGGCCAGCCCTGGCC 419

      40 ..................................................  40

     420 CTTTTGCTACCAGATCCCCACTTTTCATCTTTGTGAGAAGATCTTCTCTG 469

      41 .................................Gln.AspArgSerPr  45
                                           :::  ||||||||||
     470 CTGTCCCGGTCCTCCAGTGGGTATTTCTCTTTTGACACAGACAGGAGCCC 519

      45 oAlaProMetSerCysAspLysSerThrGlnThrProSerProProCysG 62
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     520 GGCACCCATGAGTTGTGACAAGTCAACACAAACCCCAAGTCCTCCTTGCC 569

      62 lnAlaPheAsnHisTyrLeuSerAlaMetAlaSerIleArgGlnSerGln 78
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     570 AGGCCTTCAACCACTATCTCAGTGCAATGGCTTCCATACGACAGTCTCAG 619

      79 GluGluProGluAspLeuArgProGluIleArgIleAlaGlnGluLeuAr 95
         ||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
     620 GAGGAACCTGAAGATCTGCGCCCGGAGATACGGATTGCACAGGAGCTGCG 669

      95 gArgIleGlyAspGluPheAsnGluThrTyrThrArgArgValPheAlaA 112
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     670 GCGGATCGGAGACGAGTTCAACGAAACTTACACAAGGAGGGTGTTTGCAA 719

     112 snAspTyrArgGluAlaGluAspHisProGlnMetValIleLeuGlnLeu 128
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     720 ATGATTACCGCGAGGCTGAAGACCACCCTCAAATGGTTATCTTACAACTG 769

     129 LeuArgPheIlePheArgLeuValTrpArgArgHis 140
         ||||||||||||||||||||||||||||||||||
     770 TTACGCTTTATCTTCCGTCTGGTATGGAGAAGGCAT 805

seq_name: gb_est2:BG921698

seq_documentation_block:
LOCUS       BG921698      935 bp    mRNA            EST       05-JUN-2001
DEFINITION  602825518F1 NCI_CGAP_Mam6 Mus musculus cDNA clone IMAGE:4954300 5',
            mRNA sequence.
ACCESSION   BG921698
VERSION     BG921698.1  GI:14302174
KEYWORDS    EST.
SOURCE      house mouse.
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 935)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Jeffrey Green M.D.
             cDNA Library Preparation: Life Technologies, Inc.
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Incyte Genomics, Inc.
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLAM10915  row: c  column: 05
            High quality sequence start: 3
            High quality sequence stop: 786.
FEATURES             Location/Qualifiers
     source          1. .935
                     /organism="Mus musculus"
                     /strain="FVB/N"
                     /db_xref="taxon:10090"
                     /clone="IMAGE:4954300"
                     /clone_lib="NCI_CGAP_Mam6"
                     /sex="female, virgin"
                     /tissue_type="infiltrating ductal carcinoma"
                     /dev_stage="5 months"
                     /lab_host="DH10B"
                     /note="Organ: mammary; Vector: pCMV-SPORT6; Site_1: SalI;
                     Site_2: NotI; Cloned unidirectionally.  Primer: Oligo dT.
                     Library constructed by Life Technologies. Investigator
                     providing samples: Jeffrey Green, M.D., NIH"
BASE COUNT      203 a    283 c    276 g    173 t
ORIGIN


alignment_scores:
            Quality:  583.00            Length:    197
              Ratio:  4.225               Gaps:      3
 Percent Similarity:  70.051   Percent Identity: 68.528

alignment_block:
 US-09-508-832-4 x BG921698   ..

 Align seg 1/1  to: BG921698  from: 1  to: 935

       1 MetAlaLysGlnProSerAspValSerSerGluCysAspArgGluGlyGl 17
```

```
     ||||||||||||||||||||||||||||||||||||||||||||||||||
 208 ATGGCCAAGCAACCTTCTGATGTAAGTTCTGAGTGTGACAGAGAAGGTGG 257

  17 yGlnLeuGlnProAlaGluArgProProGlnLeuArgProGlyAlaProT 34
     ||||||||||||||||||||||||||||||||||||||||||||||||||
 258 ACAATTGCAGCCTGCTGAGAGGCCTCCCCAGCTCAGGCCTGGGGCCCCTA 307

  34 hrSerLeuGlnThrGluPro..Gln..................... 41
     ||||||||||||||||||  :::
 308 CCTCCCTACAGACAGAACCGCAAGGTAATCCCGACGGCGAAGGGGACCGC 357

  41 ............................................... 41

 358 TGCCCCCACGGCATGCCCTCAGGGCCCGCTGGCCCCACCGGCCAGCCCTG 407

  41 ............................................... 41

 408 GCCCTTTTTGCTACCAGATCCCCACTTTTCATCTTTGTGAGAAGATCTTC 457

  42 .....................................AspArgS 44
                                           |||||||
 458 TCTGCTGTCCCGGTCCTCCAGTGGGTATTTCTCTTTTGACACAGACAGGA 507

  44 erProAlaProMetSerCysAspLysSerThrGlnThrProSerProPro 60
     ||||||||||||||||||||||||||||||||||  ||||||||||||
 508 GCCCGGCACCCATGAGTTGTGACAAGTCAACACAAAC.CCAAGTCCTCCT 556

  61 CysGlnAlaPheAsnHisTyrLeuSerAlaMetAlaSerIleArgGlnSe 77
     ||||||||||||||||||||||||||||||||||||||||||||||||||
 557 TGCCAGGCCTTCAACCACTATCTCAGTGCAATGGCTTCCATACGACAGTC 606

 -77 rGlnGluGluProGluAspLeuArgProGluIleArgIleAlaGlnGluL 94
     ||||||||||||||||||||||||||||||||||||||||||||||||||
 607 TCAGGAGGAACCTGAAGATCTGCGCCCGGAGATACGGATTGCACAGGAGC 656

  94 euArgArgIleGlyAspGluPheAsnGluThrTyrThrArgArgValPhe 110
     ||||||||||||||||||||||||||||||||||||||||||||||||||
 657 TGCGGCGGATCGGAGACGAGTTCAACGAAACTTACACAAGGAGGGTGTTT 706

 111 AlaAsnAspTyrArgGluAlaGluAspHisProGlnMetValIleLeuGl 127
     ||||||||||||||||||||||||  ||||||||||||||||||||||||
 707 GCAAATGATTACCGCGAGGCTGAAGA.CACCCTCAAATGGTTATCTTACA 755

 127 n..LeuLeuArgPheIlePheArgLeuValTrpArgArg 139
     |  :::|||||||||||||||||||||||||||||  :::
 756 AACGTGTTACGCTTTATCTTCCGTCTGGTATGGGCGAAA 794
```

seq_name: gb_est2:BF021882

seq_documentation_block:

LOCUS BF021882 452 bp mRNA EST 29-DEC-2000

DEFINITION uy59b09.y1 McCarrey Eddy round spermatid Mus musculus cDNA clone IMAGE:3663833 5' similar to TR:O54918 O54918 BCL2 INTERACTING MEDIATOR OF CELL DEATH ;, mRNA sequence.

ACCESSION BF021882

VERSION BF021882.1 GI:10753214

KEYWORDS EST.

SOURCE house mouse.

ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE 1 (bases 1 to 452)

AUTHORS Marra,M., Hillier,L., Kucaba,T., Martin,J., Beck,C., Wylie,T., Underwood,K., Steptoe,M., Theising,B., Allen,M., Bowers,Y., Person ,B., Swaller,T., Gibbons,M., Pape,D., Harvey,N., Schurk,R., Ritter ,E., Kohn,S., Shin,T., Jackson,Y., Cardenas,M., McCann,R., Waterston,R. and Wilson,R.

TITLE The WashU-NCI Mouse EST Project 1999

JOURNAL Unpublished (1999)

COMMENT Contact: Marra M/WashU-NCI Mouse EST Project 1999
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
Tel: 314 286 1800
Fax: 314 286 1810
Email: mouseest@watson.wustl.edu
This clone is available royalty-free through LLNL ; contact the IMAGE Consortium (info@image.llnl.gov) for further information.
MGI:1424601
Seq primer: Primer name ambiguous
High quality sequence stop: 386.

FEATURES Location/Qualifiers

source 1. .452
/organism="Mus musculus"
/strain="CD-1"
/db_xref="taxon:10090"
/clone="IMAGE:3663833"
/clone_lib="McCarrey Eddy round spermatid"
/sex="male"
/tissue_type="round spermatids, pooled from multiple mice"
/dev_stage="60 day"
/lab_host="DH10B (phage-resistant)"
/note="Organ: testis; Vector: pBluescript SK+ (Stratagene); Site_1: XhoII; Site_2: EcoRI; cDNA oligo dT-primed [5'-(GA)10-ACTAGTCTCGAGTTTTTTTTTTTTT-3'] and directionally cloned using 5' linkers 5'-AATTCGGCACGAG-3' and 5'-CTCGTGCCG-3'. Size selection of >400bp material gives average insert size ranging from 1-2 kb. Library was mass excised (from lambda-UniZAP-XR) and resulting single-stranded phagemids were prepped and tranformed into DH10B. Library contains 98.5% recombinants. References: J. Androl. 20:635-639 and Gene 25:263-269. Library constructed and donated by J. McCarrey, Ph.D. (Southwest Foundation for Biomedical Research, Dept. of Genetics); excision done by E.M. Eddy, Ph.D. (National Institutes of Health, National Institute of Environmental Health Sciences). Original lambda-based library is available through ATCC, catalog #63423."

BASE COUNT 106 a 130 c 112 g 104 t

ORIGIN

alignment_scores:
Quality: 525.00 Length: 99
Ratio: 5.303 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
US-09-508-832-4 x BF021882 ..

Align seg 1/1 to: BF021882 from: 1 to: 452

```
  42 AspArgSerProAlaProMetSerCysAspLysSerThrGlnThrProSe 58
     ||||||||||||||||||||||||||||||||||||||||||||||||||
 121 GACAGGAGCCCGGCACCCATGAGTTGTGACAAGTCAACACAAACCCCAAG 170

  58 rProProCysGlnAlaPheAsnHisTyrLeuSerAlaMetAlaSerIleA 75
     ||||||||||||||||||||||||||||||||||||||||||||||||||
 171 TCCTCCTTGCCAGGCCTTCAACCACTATCTCAGTGCAATGGCTTCCATAC 220

  75 rgGlnSerGlnGluGluProGluAspLeuArgProGluIleArgIleAla 91
     ||||||||||||||||||||||||||||||||||||||||||||||||||
 221 GACAGTCTCAGGAGGAACCTGAAGATCTGCGCCCGGAGATACGGATTGCA 270

  92 GlnGluLeuArgArgIleGlyAspGluPheAsnGluThrTyrThrArgAr 108
     ||||||||||||||||||||||||||||||||||||||||||||||||||
 271 CAGGAGCTGCGGCGGATCGGAGACGAGTTCAACGAAACTTACACAAGGAG 320

 108 gValPheAlaAsnAspTyrArgGluAlaGluAspHisProGlnMetValI 125
     ||||||||||||||||||||||||||||||||||||||||||||||||||
 321 GGTGTTTGCAAATGATTACCGCGAGGCTGAAGACCACCCTCAAATGGTTA 370

 125 leLeuGlnLeuLeuArgPheIlePheArgLeuValTrpArgArgHis 140
     ||||||||||||||||||||||||||||||||||||||||||||||||||
 371 TCTTACAACTGTTACGCTTTATCTTCCGTCTGGTATGGAGAAGGCAT 417
```

seq_name: gb_est1:AI971169

```
seq_documentation_block:
LOCUS       AI971169      492 bp    mRNA            EST       08-MAR-2000
DEFINITION  wr24h12.x1 NCI_CGAP_Pr28 Homo sapiens cDNA clone IMAGE:2488679 3'
            similar to TR:O43522 O43522 BIML. [1] ;, mRNA sequence.
ACCESSION   AI971169
VERSION     AI971169.1  GI:5767995
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 492)
  AUTHORS   NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
  TITLE     National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
            Tumor Gene Index
  JOURNAL   Unpublished (1997)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Michael J. Brownstein, M.D., Ph.D., Michael R.
            Emmert-Buck, M.D., Ph.D.
             cDNA Library Preparation: M. Bento Soares, Ph.D.
             cDNA Library Arrayed by: Greg Lennon, Ph.D.
             DNA Sequencing by: Washington University Genome Sequencing Center
             Clone distribution: NCI-CGAP clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            www-bio.llnl.gov/bbrp/image/image.html
            Insert Length: 712   Std Error: 0.00
            Seq primer: -40UP from Gibco
            High quality sequence stop: 450.
FEATURES             Location/Qualifiers
     source          1. .492
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:2488679"
                     /clone_lib="NCI_CGAP_Pr28"
                     /sex="male"
                     /dev_stage="adult"
                     /lab_host="DH10B"
                     /note="Organ: prostate; Vector: pT7T3D-Pac (Pharmacia)
                     with a modified polylinker; Plasmid DNA from the
                     normalized library NCI_CGAP_Pr22 was prepared, and ss
                     circles were made in vitro. Following HAP purification,
                     this DNA was used as tracer in a subtractive hybridization
                     reaction. The driver was PCR-amplified cDNAs from a pool
                     of 5,000 clones made from the same library (cloneIDs
                     985608-986759, 1101192-1101959, and 1217928-1220615).
                     Subtraction by Bento Soares and M. Fatima Bonaldo.  "
BASE COUNT      119 a    107 c    130 g    134 t      2 others
ORIGIN


alignment_scores:
            Quality:  370.00            Length:     74
              Ratio:  5.211               Gaps:      0
 Percent Similarity:  95.946   Percent Identity: 94.595

alignment_block:
 US-09-508-832-4 x AI971169/rev   ..

 Align seg 1/1  to reverse of: AI971169  from: 1  to: 492

       1 MetAlaLysGlnProSerAspValSerSerGluCysAspArgGluGlyGl 17
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     425 ATGGCAAAGCAACCTTCTGATGTAAGTTCTGAGTGTGACCGAGAAGGTAG 376

      17 yGlnLeuGlnProAlaGluArgProProGlnLeuArgProGlyAlaProT 34
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     375 ACAATTGCAGCCTGCGGAGAGGCCTCCCCAGCTCAGACCTGGGGCCCCTA 326

      34 hrSerLeuGlnThrGluProGlnAspArgSerProAlaProMetSerCys 50
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     325 CCTCCCTACAGACAGAGCCACAAGACAGGAGCCCAGCACCCATGAGTTGT 276

      51 AspLysSerThrGlnThrProSerProProCysGlnAlaPheAsnHisTy 67
         ||||||||||||||||||   ||||||||||||||||||||||||||||||
     275 GACAAATCAACACAAACNNCAAGTCCTCCTTGCCAGGCCTTCAACCACTA 226

      67 rLeuSerAlaMetAlaSerIle 74
         |||||||||||||:::   |||
     225 TCTCAGTGCAATGGTAGTCATC 204

seq_name: gb_est2:BG173095

seq_documentation_block:
LOCUS       BG173095      668 bp    mRNA            EST       06-FEB-2001
DEFINITION  602336666F1 NCI_CGAP_Mam1 Mus musculus cDNA clone IMAGE:4459720 5',
            mRNA sequence.
ACCESSION   BG173095
VERSION     BG173095.1  GI:12679707
KEYWORDS    EST.
SOURCE      house mouse.
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 668)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Gilbert Smith, Ph.D.
             cDNA Library Preparation: Life Technologies, Inc.
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Incyte Genomics, Inc.
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLAM10260  row: c  column: 17
            High quality sequence stop: 599.
FEATURES             Location/Qualifiers
     source          1. .668
                     /organism="Mus musculus"
                     /strain="FVB/N"
                     /db_xref="taxon:10090"
                     /clone="IMAGE:4459720"
                     /clone_lib="NCI_CGAP_Mam1"
                     /tissue_type="tumor, biopsy sample"
                     /dev_stage="3 months, virgin"
                     /lab_host="DH10B"
                     /note="Organ: mammary; Vector: pCMV-SPORT6; Site_1: SalI;
                     Site_2: NotI; Cloned unidirectionally.  Primer: Oligo dT.
                     Library constructed by Life Technologies. Investigator
                     providing samples: Gilbert Smith, NIH"
BASE COUNT      135 a    235 c    162 g    136 t
ORIGIN


alignment_scores:
            Quality:  354.00            Length:    148
              Ratio:  4.265               Gaps:      3
 Percent Similarity:  56.081   Percent Identity: 54.054

alignment_block:
 US-09-508-832-4 x BG173095   ..

 Align seg 1/1  to: BG173095  from: 1  to: 668

       1 MetAlaLysGlnProSerAspValSerSerGluCysAspArgGluGlyGl 17
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     223 ATGGCCAAGCAACCTTCTGATGTAAGTTCTGAGTGTGACAGAGAAGGTGG 272

      17 yGlnLeuGlnProAlaGluArgProProGlnLeuArgProGlyAlaProT 34
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     273 ACAATTGCAGCCTGCTGAGAGGCCTCCCCAGCTCAGGCCTGGGGCCCCTA 322
```

```
      34 hrSerLeuGlnThrGluPro..............................  40
         ||||||||||||||||||||
     323 CCTCCCTACAGACAGAACCGCAAGGTAATCCCGACGGCGAAGGGGACCTG 372

      40 ..................................................  40

     373 CTGCCCCCACGGCAGCCCTCAGGGCCCGCTTGGCCCCACCGGCAGCCCTG 422

      40 ..................................................  40

     423 GCCCTTTTGCTACCAGATCCCCCACTTTTCATCTTTGTGAGAAGATCTTC 472

      41 .............................................GlnAsp  42
                                                        |||
     473 TCTGCTGTCCCGGTCCTCCCAGTGGGTATATTCTCCTTTTGACACAGCAC 522

      43 ArgSerProAlaProMetSerCysAspLysSerThrGlnThrProSerPr  59
         ||||||||||||||||||||      ::::::   :::||||||||||||
     523 AGGAGCCCGGCACCCATGAGTGTGACAAGTCAACACAAAACCCCAAGTCC 572

      59 oProCysGlnAlaPheAsnHisTyrLeuSer.AlaMetAlaSerIleArg  75
         |||||||||||||||||||||||||||||| |||||||||   ||||||
     573 TCCTTGCCAGGCCTTCAACCACTATCTCAGTTGCAATGGCTTTCATACGA 622

      76 GlnSer.GlnGluGluProGluAspLeuArgProGluIle 88
         |||||| ||||||||||||||||||||||||||||||||
     623 CAGTCTCCAGGAGGAACCTGAGGATCTGCGCCCGGAGATC 662

seq_name: gb_gss:AZ706148

seq_documentation_block:
LOCUS       AZ706148      580 bp    DNA             GSS       24-JAN-2001
DEFINITION  RPCI-23-227P3.TV RPCI-23 Mus musculus genomic clone RPCI-23-227P3,
            DNA sequence.
ACCESSION   AZ706148
VERSION     AZ706148.1  GI:12433319
KEYWORDS    GSS.
SOURCE      house mouse.
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 580)
  AUTHORS   Zhao,S., Nierman,W., Feldblyum,T., Malek,J., Shatsman,S., Akinret
            ,B., Levins,M., Mcgann,S., Tsegaye,G., Geer,K., Krol,M., de Jong,P.
            and Fraser,C.M.
  TITLE     Mouse BAC End Sequences from Library RPCI-23
  JOURNAL   Unpublished (1999)
COMMENT     Other_GSSs: RPCI-23-227P3.TJ
            Contact: Shaying Zhao
            Department of Eukaryotic Genomics
            The Institute for Genomic Research
            9712 Medical Center Dr., Rockville, MD 20850, USA
            Tel: 301 838 0200
            Fax: 301 838 0208
            Email: szhao@tigr.org
            Clones are derived from the mouse BAC library RPCI-23. For BAC
            library availability, please contact Pieter de Jong
            (pdejong@mail.cho.org). Clones may be purchased from BACPAC
            Resources (http://www.chori.org/bacpac/orderingframe.htm). BAC end
            page: http://ww .tigr.org/tdb/bac_ends/mouse/bac_end_intro.html
            Plate: 227  row: P  column: 3
            Seq primer: T7
            Class: BAC ends.
FEATURES             Location/Qualifiers
     source          1. .580
                     /organism="Mus musculus"
                     /strain="C57BL/6J"
                     /db_xref="taxon:10090"
                     /clone="RPCI-23-227P3"
                     /clone_lib="RPCI-23"
                     /sex="Female"
                     /lab_host="DH10B"
                     /note="Organ: Kidney/Brain; Vector: pBACe3.6; Site_1:
                     EcoRI; Site_2: EcoRI; Female C57BL/6J mouse kidney and/or
                     brain genomic DNA was isolated and  partially digested
                     with a combination of EcoRI and EcoRI Methylase.   Size
                     selected DNA was cloned into the pBACe3.6 vector at the
                     EcoRI  sites.  The ligation products were transformed into
                     DH10B  electrocompetent cells (BRL Life Technologies).  "
BASE COUNT      138 a    162 c    138 g    142 t
ORIGIN

alignment_scores:
            Quality:  347.00            Length:    127
              Ratio:  4.887               Gaps:      1
 Percent Similarity:  55.906   Percent Identity: 55.906

alignment_block:
US-09-508-832-4 x AZ706148   ..

Align seg 1/1  to: AZ706148  from: 1  to: 580

       1 MetAlaLysGlnProSerAspValSerSerGluCysAspArgGluGlyGl 17
         ||||||||||||||||||||||||||||||||||||||||||||||||||
      90 ATGGCCAAGCAACCTTCTGATGTAAGTTCTGAGTGTGACAGAGAAGGTGG 139

      17 yGlnLeuGlnProAlaGluArgProProGlnLeuArgProGlyAlaProT 34
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     140 ACAATTGCAGCCTGCTGAGAGGCCTCCCCAGCTCAGGCCTGGGGCCCCTA 189

      34 hrSerLeuGlnThrGluProGln.............................  41
         |||||||||||||||||||||||
     190 CCTCCCTACAGACAGAACCGCAAGGTAATCCCGACGGCGAAGGGGACCGC 239

      41 ..................................................  41

     240 TGCCCCCACGGCAGCCCTCAGGGCCCGCTGGCCCCACCGGCCAGCCCTGG 289

      41 ..................................................  41

     290 CCCTTTTGCTACCAGATCCCCACTTTTCATCTTTGTGAGAAGATCTTCTC 339

      42 .............................................AspArgSer  44
                                                        |||||||||
     340 TGCTGTCCCGGTCCTCCAGTGGGTATTTCTCTTTTGACACAGACAGGAGC 389

      45 ProAlaProMetSerCysAspLysSerThrGlnThrProSerProProCy  61
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     390 CCGGCACCCATGAGTTGTGACAAGTCAACACAAACCCCAAGTCCTCCTTG 439

      61 sGlnAlaPheAsnHisTyrLeuSerAlaMet 71
         |||||||||||||||||||||||||||||||
     440 CCAGGCCTTCAACCACTATCTCAGTGCAATG 470

seq_name: gb_est1:AW629314

seq_documentation_block:
LOCUS       AW629314      664 bp    mRNA            EST       31-MAR-2000
DEFINITION  hi56e02.x1 Soares_NFL_T_GBC_S1 Homo sapiens cDNA clone
            IMAGE:2976314 3' similar to TR:O43522 O43522 BIML. [1] ;, mRNA
            sequence.
ACCESSION   AW629314
VERSION     AW629314.1  GI:7376104
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 664)
  AUTHORS   NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
  TITLE     National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
            Tumor Gene Index
  JOURNAL   Unpublished (1997)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
```

```
            This clone is available royalty-free through LLNL ; contact the
            IMAGE Consortium (info@image.llnl.gov) for further information.
            Seq primer: -40UP from Gibco
            High quality sequence stop: 458.
FEATURES             Location/Qualifiers
     source          1. .664
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:2976314"
                     /clone_lib="Soares_NFL_T_GBC_S1"
                     /lab_host="DH10B"
                     /note="Organ: pooled; Vector: pT7T3D-Pac (Pharmacia) with
                     a modified polylinker; Site_1: Not I; Site_2: Eco RI;
                     Equal amounts of plasmid DNA from three normalized
                     libraries (fetal lung NbHL19W, testis NHT, and B-cell
                     NCI_CGAP_GCB1) were mixed, and ss circles were made in
                     vitro. Following HAP purification, this DNA was used as
                     tracer in a subtractive hybridization reaction. The driver
                     was PCR-amplified cDNAs from pools of 5,000 clones made
                     from the same 3 libraries. The pools consisted of
                     I.M.A.G.E. clones 297480-302087, 682632-687239,
                     726408-728711, and 729096-731399. Subtraction by Bento
                     Soares and M. Fatima Bonaldo. "
BASE COUNT      176 a    131 c    148 g    208 t      1 others
ORIGIN


alignment_scores:
            Quality:  287.00            Length:     69
              Ratio:  4.864               Gaps:      0
 Percent Similarity:  85.507   Percent Identity: 78.261

alignment_block:
 US-09-508-832-4 x AW629314/rev   ..

 Align seg 1/1  to reverse of: AW629314  from: 1  to: 664

       3 LysGlnProSerAspValSerSerGluCysAspArgGluGlyGlyGlnLe 19
         |||   |||   |||   |||   ||||||      ||||||   |||::
     664 AAGCCACCTTTTGATGAAAGTTTTGAGTGTGCCNGAGAAGGTAGACAAAT 615

      19 uGlnProAlaGluArgProProGlnLeuArgProGlyAlaProThrSerL 36
         :||||||||||:::|||   ||||||:::|||   ||||||||||||||
     614 GCAGCCTGCGAAGAGGCTTCCCCAGTTCAGACTTGGGGCCCCTACCTCCT 565

      36 euGlnThrGluProGlnAspArgSerProAlaProMetSerCysAspLys 52
           :::|||||||||||||||||||||||||||||||||||||||||||||
     564 ACAAGACAGAGCCACAAGACAGGAGCCCAGCACCCATGAGTTGTGACAAA 515

      53 SerThrGlnThrProSerProProCysGlnAlaPheAsnHisTyrLeuSe 69
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     514 TCAACACAAACCCCAAGTCCTCCTTGCCAGGCCTTCAACCACTATCTCAG 465

      69 rAlaMet 71
         |:::|||
     464 TGTAATG 458

seq_name: gb_est2:BF319454

seq_documentation_block:
LOCUS       BF319454      389 bp    mRNA            EST       29-DEC-2000
DEFINITION  uy59b09.x1 McCarrey Eddy round spermatid Mus musculus cDNA clone
            IMAGE:3663833 3' similar to TR:O54918 O54918 BCL2 INTERACTING
            MEDIATOR OF CELL DEATH ;, mRNA sequence.
ACCESSION   BF319454
VERSION     BF319454.1  GI:11268195
KEYWORDS    EST.
SOURCE      house mouse.
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 389)
  AUTHORS   Marra,M., Hillier,L., Kucaba,T., Martin,J., Beck,C., Wylie,T.,
            Underwood,K., Steptoe,M., Theising,B., Allen,M., Bowers,Y., Person
            ,B., Swaller,T., Gibbons,M., Pape,D., Harvey,N., Schurk,R., Ritter
            ,E., Kohn,S., Shin,T., Jackson,Y., Cardenas,M., McCann,R.,
            Waterston,R. and Wilson,R.
  TITLE     The WashU-NCI Mouse EST Project 1999
  JOURNAL   Unpublished (1999)
COMMENT     Other_ESTs: uy59b09.y1
            Contact: Marra M/WashU-NCI Mouse EST Project 1999
            Washington University School of Medicine
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: mouseest@watson.wustl.edu
            This clone is available royalty-free through LLNL ; contact the
            IMAGE Consortium (info@image.llnl.gov) for further information.
            MGI:1424601
            High quality sequence stop: 325.
FEATURES             Location/Qualifiers
     source          1. .389
                     /organism="Mus musculus"
                     /strain="CD-1"
                     /db_xref="taxon:10090"
                     /clone="IMAGE:3663833"
                     /clone_lib="McCarrey Eddy round spermatid"
                     /sex="male"
                     /tissue_type="round spermatids, pooled from multiple mice"
                     /dev_stage="60 day"
                     /lab_host="DH10B (phage-resistant)"
                     /note="Organ: testis; Vector: pBluescript SK+ (Stratagene
                     ); Site_1: XhoII; Site_2: EcoRI; cDNA oligo dT-primed
                     [5'-(GA)10-ACTAGTCTCGAGTTTTTTTTTTTT-3'] and directionally
                     cloned  using 5' linkers 5'-AATTCGGCACGAG-3' and
                     5'-CTCGTGCCG-3'. Size selection of >400bp material gives
                     average insert size ranging from 1-2 kb. Library was mass
                     excised (from  lambda-UniZAP-XR) and resulting
                     single-stranded phagemids were prepped and tranformed
                     into DH10B. Library contains 98.5% recombinants.
                     References: J. Androl. 20:635-639 and Gene 25:263-269.
                     Library constructed and donated by J. McCarrey, Ph.D.
                     (Southwest Foundation for Biomedical Research, Dept. of
                     Genetics); excision done by E.M. Eddy, Ph.D. (National
                     Institutes of  Health, National Institute of Environmental
                     Health Sciences).  Original lambda-based library is
                     available through ATCC, catalog #63423."
BASE COUNT      100 a    104 c     89 g     96 t
ORIGIN


alignment_scores:
            Quality:  249.00            Length:     53
              Ratio:  4.882               Gaps:      1
 Percent Similarity:  96.226   Percent Identity: 94.340

alignment_block:
 US-09-508-832-4 x BF319454/rev   ..

 Align seg 1/1  to reverse of: BF319454  from: 1  to: 389

      89 ArgIleAlaGlnGluLeuArgArgIleGly.AspGluPheAsnGluThrT 105
         |||:::|||||||||||||||   |||||| ||||||||||||||||||
     389 CGATTTGCACAGGAACTTCGGGGGATCGGAAGACGAGTTCAACGAAACTT 340

     105 yrThrArgArgValPheAlaAsnAspTyrArgGluAlaGluAspHisPro 121
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     339 ACACAAGGAGGGTGTTTGCAAATGATTACCGCGAGGCTGAAGACCACCCT 290

     122 GlnMetValIleLeuGlnLeuLeuArgPheIlePheArgLeuValTrpAr 138
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     289 CAAATGGTTATCTTACAACTGTTACGCTTTATCTTCCGTCTGGTATGGAG 240

     138 gArgHis 140
         |||||||
     239 AAGGCAT 233
```

```
seq_name: gb_est1:AA629308

seq_documentation_block:
LOCUS       AA629308      501 bp    mRNA            EST       16-OCT-1997
DEFINITION  zu84g06.s1 Soares_testis_NHT Homo sapiens cDNA clone IMAGE:744730
            3', mRNA sequence.
ACCESSION   AA629308
VERSION     AA629308.1  GI:2541695
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 501)
  AUTHORS   Hillier,L., Allen,M., Bowles,L., Dubuque,T., Geisel,G., Jost,S.,
            Krizman,D., Kucaba,T., Lacy,M., Le,N., Lennon,G., Marra,M., Martin
            ,J., Moore,B., Schellenberg,K., Steptoe,M., Tan,F., Theising,B.,
            White,Y., Wylie,T., Waterston,R. and Wilson,R.
  TITLE     WashU-NCI human EST Project
  JOURNAL   Unpublished (1997)
COMMENT     Contact: Wilson RK
            Washington University School of Medicine
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: est@watson.wustl.edu
            This clone is available royalty-free through LLNL ; contact the
            IMAGE Consortium (info@image.llnl.gov) for further information.
            Seq primer: -40m13 fwd. ET from Amersham
            High quality sequence stop: 471.
FEATURES             Location/Qualifiers
     source          1. .501
                     /organism="Homo sapiens"
                     /db_xref="GDB:5932418"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:744730"
                     /clone_lib="Soares_testis_NHT"
                     /sex="male"
                     /lab_host="DH10B"
                     /note="Vector: pT7T3D-Pac (Pharmacia) with a modified
                     polylinker; Site_1: Not I; Site_2: Eco RI; 1st strand cDNA
                     was prepared from mRNA obtained from Clontech Laboratories
                     , Inc., and primed with  a Not I - oligo(dT) primer [5'
                     TGTTACCAATCTGAAGTGGGAGCGGCCGCCCAATTTTTTTTTTTTTTTT 3']. 
                     Double-stranded cDNA was ligated to Eco RI adaptors
                     (Pharmacia), digested with Not I and cloned into the Not I
                     and Eco RI sites of the modified pT7T3 vector. Library
                     went through one round of normalization to Cot5, and was
                     constructed  by Bento Soares and M. Fatima Bonaldo. "
BASE COUNT      155 a    112 c     97 g    137 t
ORIGIN


alignment_scores:
           Quality:  127.00            Length:     32
             Ratio:  4.536               Gaps:      0
 Percent Similarity:  87.500   Percent Identity: 71.875

alignment_block:
 US-09-508-832-4 x AA629308/rev   ..

 Align seg 1/1  to reverse of: AA629308  from: 1  to: 501

     109 ValPheAlaAsnAspTyrArgGluAlaGluAspHisProGlnMetValIl 125
         ||||||   |||:::|||:::   |||||||||||||||:::||||||||
     501 GTATTTTTGAATAATTACCAAGCAGCCGAAGACCACCCACGAATGGTTAT 452

     125 eLeuGlnLeuLeuArgPheIlePheArgLeuValTrpArgArgHis 140
         ||||:::|||||||||:::|||   |||||||||||||||   |||
     451 CTTACGACTGTTACGTTACATTGTCCGCCTGGTGTGGAGAATGCAT 406

seq_name: gb_est2:BF172831

seq_documentation_block:
LOCUS       BF172831      210 bp    mRNA            EST       23-MAR-2001
DEFINITION  PCL5805 Myeloma (PCL) cDNA library Homo sapiens cDNA, mRNA
            sequence.
ACCESSION   BF172831
VERSION     BF172831.1  GI:13439045
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 210)
  AUTHORS   Claudio,J.O., Tang,H., Khan,E.M., Voralia,M., Li,Z., Cukerman,E.,
            Franciso-Pabalan,O., Liew,C.C. and Stewart,A.K.
  TITLE     The transcriptional phenotype of myeloma cells
  JOURNAL   Unpublished (2000)
COMMENT     Contact: A. Keith Stewart, M.D.
            Oncology Research
            University Health Network
            610 University Ave., 5-126, Toronto, Ontario, M5G 2M9, Canada
            Tel: (416) 946-4639
            Fax: (416) 946-6546
            Email: k.stewart@utoronto.ca
            PCR PRimers
            FORWARD: 5'-GCCAAGCTCGAAATTAACCCTCACTAAAGGG-3'
            BACKWARD: 5'-CCAGTGAATTGTAATACGACTCACTATAGGGCG-3'
            Seq primer: 5'-GAAATTAACCCTCACTAAAGG-3'.
FEATURES             Location/Qualifiers
     source          1. .210
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone_lib="Myeloma (PCL) cDNA library"
                     /sex="male"
                     /tissue_type="Blood"
                     /cell_type="myeloma"
                     /dev_stage="Plasma cell leukemia"
                     /note="Vector: Lambda Zap Express; Site_1: EcoRI; Site_2:
                     XhoI; mRNA was purified from plasma cell leukemia
                     patient's peripheral blood  containing >95% myeloma. An
                     oligo d(T)18 primer containing XhoI restriction  site was
                     used to prime first strand synthesis using M-MLV reverse
                     transcriptase. To protect the cDNAs from XhoI digestion in
                     subsequent  cloning step, the nucleotide analogue
                     5-methyl-dCTP was added to the  nucleotide mixture and
                     [a-32P]dATP was added to monitor the quantity and  quality
                     of first strand synthesis. After second-strand synthesis
                     and  blunting of cDNA termini, EcoRI adapters were ligated
                     , followed by kinase  treatment and digestion with XhoI.
                     The cDNAs were then size-fractionated  using Sephacryl
                     S-500 column and then ligated into EcoRI and XhoI digested
                     Lambda Zap Express vector. The ligation product was
                     packaged using Gigapack  II packaging extract. The library
                     had primary titre of approx. 1x106. Clones from  the
                     primary library were randomly selected for single pass
                     sequencing."
BASE COUNT       49 a     40 c     57 g     59 t      5 others
ORIGIN


alignment_scores:
           Quality:  126.00            Length:     39
             Ratio:  3.818               Gaps:      2
 Percent Similarity:  84.615   Percent Identity: 76.923

alignment_block:
 US-09-508-832-4 x BF172831   ..

 Align seg 1/1  to: BF172831  from: 1  to: 210

      72 AlaSerIleArgGlnSerGlnGluGluProGluAspLeuArgProGluIl 88
         ||||||:::|||||||||      ||||||   |||:::||||||||||||
      95 GCTTCCATGAGGCAGTCT......GAACCTGCAGATATGCGCCCAGAGAT 138
```

```
     88 eArgIleAlaGlnGluLeu.ArgArgIleGlyAspGluPheAsnGluThr 104
        |   |||||||||||||||| |||||||||||||||||||||||||   :::
    139 ATGGATCGCCCAAGAGTTGGCGGCGTATTGGAGACGAGTTTAACGCTAAC 188

    105 TyrThrArgArgVal 109
        ||||||||||||||
    189 TATACAAGGAGGGTA 203
```

seq_name: gb_est1:AA629050

seq_documentation_block:

```
LOCUS       AA629050      501 bp    mRNA            EST       16-OCT-1997
DEFINITION  zu84a06.s1 Soares_testis_NHT Homo sapiens cDNA clone IMAGE:744658
            3', mRNA sequence.
ACCESSION   AA629050
VERSION     AA629050.1  GI:2541437
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 501)
  AUTHORS   Hillier,L., Allen,M., Bowles,L., Dubuque,T., Geisel,G., Jost,S.,
            Krizman,D., Kucaba,T., Lacy,M., Le,N., Lennon,G., Marra,M., Martin
            ,J., Moore,B., Schellenberg,K., Steptoe,M., Tan,F., Theising,B.,
            White,Y., Wylie,T., Waterston,R. and Wilson,R.
  TITLE     WashU-NCI human EST Project
  JOURNAL   Unpublished (1997)
COMMENT     Contact: Wilson RK
            Washington University School of Medicine
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: est@watson.wustl.edu
            This clone is available royalty-free through LLNL ; contact the
            IMAGE Consortium (info@image.llnl.gov) for further information.
            Seq primer: -40m13 fwd. ET from Amersham
            High quality sequence stop: 474.
FEATURES             Location/Qualifiers
     source          1. .501
                     /organism="Homo sapiens"
                     /db_xref="GDB:5932318"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:744658"
                     /clone_lib="Soares_testis_NHT"
                     /sex="male"
                     /lab_host="DH10B"
                     /note="Vector: pT7T3D-Pac (Pharmacia) with a modified
                     polylinker; Site_1: Not I; Site_2: Eco RI; 1st strand cDNA
                     was prepared from mRNA obtained from Clontech Laboratories
                     , Inc., and primed with  a Not I - oligo(dT) primer [5'
                     TGTTACCAATCTGAAGTGGGAGCGGCCGCCCAATTTTTTTTTTTTTTTT 3'].
                     Double-stranded cDNA was ligated to Eco RI adaptors
                     (Pharmacia), digested with Not I and cloned into the Not I
                     and Eco RI sites of the modified pT7T3 vector. Library
                     went through one round of normalization to Cot5, and was
                     constructed  by Bento Soares and M. Fatima Bonaldo. "
BASE COUNT      155 a    111 c     98 g    137 t
ORIGIN
```

alignment_scores:

```
           Quality:  123.00            Length:     31
             Ratio:  4.556               Gaps:      0
 Percent Similarity:  87.097   Percent Identity: 70.968
```

alignment_block:

US-09-508-832-4 x AA629050/rev ..

Align seg 1/1 to reverse of: AA629050 from: 1 to: 501

```
    110 PheAlaAsnAspTyrArgGluAlaGluAspHisProGlnMetValIleLe 126
```

```
        |||   |||:::|||:::   |||||||||||||||:::||||||||||
    499 TTTTTGAATAATTACCAAGCAGCCGAAGACCACCCACGAATGGTTATCTT 450

    126 uGlnLeuLeuArgPheIlePheArgLeuValTrpArgArgHis 140
        |:::|||||||||:::|||   ||||||||||||||   |||
    449 ACGACTGTTACGTTACATTGTCCGCCTGGTGTGGAGAATGCAT 407
```

seq_name: gb_est1:AF209718

seq_documentation_block:

```
LOCUS       AF209718      537 bp    mRNA            EST       30-MAY-2000
DEFINITION  AF209718 Xenopus laevis intestine adult Xenopus laevis cDNA clone
            pXLG10 similar to Mus musculus BimEL, mRNA sequence.
ACCESSION   AF209718
VERSION     AF209718.1  GI:8110110
KEYWORDS    EST.
SOURCE      African clawed frog.
  ORGANISM  Xenopus laevis
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidea; Pipidae;
            Xenopodinae; Xenopus.
REFERENCE   1  (bases 1 to 537)
  AUTHORS   Minter,R.
  TITLE     Development of Antibody Technology to Identify Natural Killer Cell
            Surface Antigens in Xenopus laevis. Thesis (1999) University of
            Durham, South Rd., Durham, UK
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Minter,R., Horton,J.D. and Watson,M.D.
            Biological Sciences
            University of Durham
            South Road, Durham, DH1 3LE, UK
            Email: martin.watson@durham.ac.uk.
FEATURES             Location/Qualifiers
     source          1. .537
                     /organism="Xenopus laevis"
                     /db_xref="taxon:8355"
                     /clone="pXLG10"
                     /clone_lib="Xenopus laevis intestine adult"
                     /tissue_type="intestine"
                     /cell_type="epithelial lymphocyte"
                     /dev_stage="adult"
BASE COUNT      126 a    129 c    143 g    139 t
ORIGIN
```

alignment_scores:

```
           Quality:   98.50            Length:     64
             Ratio:  2.239               Gaps:      3
 Percent Similarity:  68.750   Percent Identity: 45.312
```

alignment_block:

US-09-508-832-4 x AF209718 ..

Align seg 1/1 to: AF209718 from: 1 to: 537

```
      1 MetAlaLysGlnProSerAspValSerSerGluCysAspArg...GluGl 16
        |||||||||||||||||   :::|||   ||||||:::      |||||
    223 ATGGCCAAACAACCGTCGGTCTTGAGTCCGGAGTGTAATAGTGGTGAAGG 272

     16 yGlyGlnLeuGlnProAlaGluArgProProGlnLeuArgPro..... 30
        ||||||||||||   ::::::|||      :::   ||||||
    273 TGGCCAGTTACAATCAACAAGCAGACGACATTCTCATCGTCCTCGCAGAA 322

     31 ..GlyAlaProThrSerLeuGlnThrGluProGlnAspArgSerProAla 46
          |||||||||||||||||   ::::::|||   :::   :::   :::
    323 GAGGGGCCCCCACCTCTCTT...AGCAGTCCTTTTCAAGGTAATCAATCA 369

     47 ProMetSerCysAspLysSerThrGlnThrProSerProProPro 60
           ::::   ::::::|||::::::||||||      |||
    370 GATGAGGGTGGGAGCTCCTCAGCCAGCACTCCTTGGGGTCCT 411
```

seq_name: gb_est2:BG649655

```
seq_documentation_block:
LOCUS       BG649655      383 bp    mRNA            EST       24-APR-2001
DEFINITION  RHIZ2_82_D01.b1_A003 Rhizome2 (RHIZ2) Sorghum propinquum cDNA, mRNA
            sequence.
ACCESSION   BG649655
VERSION     BG649655.1  GI:13784767
KEYWORDS    EST.
SOURCE      Sorghum propinquum.
  ORGANISM  Sorghum propinquum
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC
            clade; Panicoideae; Andropogoneae; Sorghum.
REFERENCE   1  (bases 1 to 383)
  AUTHORS   Cordonnier-Pratt,M.-M., Gingle,A., Paterson,A., Sudman,M. and Pratt
            ,L.H.
  TITLE     An EST database from Sorghum: Sorghum propinquum rhizomes
  JOURNAL   Unpublished (2000)
COMMENT     Contact: Cordonnier-Pratt MM
            Department of Botany
            The University of Georgia
            Plant Sciences Building, Rm. 2502, Athens, GA 30602-7271, USA
            Tel: 706 542 1860
            Fax: 706 542 1805
            Email: mmpratt@uga.edu
            Seq primer: JEN REV
            High quality sequence stop: 300
            POLYA=No.
FEATURES             Location/Qualifiers
     source          1. .383
                     /organism="Sorghum propinquum"
                     /db_xref="taxon:132711"
                     /clone_lib="Rhizome2 (RHIZ2)"
                     /note="Organ: Rhizomes; Vector: pBluescript II from Lambda
                     Zap II; Site_1: XhoI; Site_2: EcoRI; The library was made
                     from poly-A RNA in the cloning vector lambda ZAP II.
                     Clones to be sequenced were prepared by mass excision."
BASE COUNT       52 a    148 c    126 g     57 t
ORIGIN


alignment_scores:
            Quality:   98.00            Length:     99
              Ratio:  1.815              Gaps:      5
 Percent Similarity:  54.545   Percent Identity: 34.343

alignment_block:
 US-09-508-832-4 x BG649655   ..

 Align seg 1/1  to: BG649655  from: 1  to: 383

       2 AlaLysGlnProSerAspValSerSerGluCysAspArgGlu........ 15
         :::||||||||||         :::||||||||||:::|||
      48 TCGAAACAGCCCCCT......GCGTCCGAATGTTCTAGAGCCTTCCGGCC 91

      16 .......GlyGlyGlnLeuGlnProAlaGlu.ArgProProGlnLeuArg 29
                 |||      :::   |||   ||| |||   |||      |||
      92 GGGCCTAGGGTTCCCGATGGGGCCCGAGGAGGCGGAGGCCGTGCTGGAGA 141

      30 Pro...GlyAlaProThrSerLeuGlnThrGluProGlnAspArgSerPr 45
         |||   |||:::   |||:::   ::::::   |||:::      |||||
     142 CCATCTGGGACCTCCACGACAAGGTCAGCGACGCCATCCACGCCCTCTCC 191

      45 oAlaProMetSerCysAspLysSerThrGlnThrPro.............S 58
         |||||||   ||||||      |||:::   :::|||             |
     192 CGCGCCCACTTCCTGCGCGCCGTCCGCCGCCGCGCCTCCGCCGCGGACAA 241

      58 erProProCysGlnAlaPheAsnHisTyrLeuSerAlaMetAlaSerIle 74
         ||||||||      :::   :::         ::::::   ||||||
     242 GCCCGCCGCCGGCCTCGTGTACGTCAAGGGCGGGGGTCTGGGCCTCGGCG 291

      75 ArgGlnSerGlnGluGluProGluAspLeuArgProGluIleArg 89
         |||   ::::::   :::            ||||||::::::|||
     292 CGCGCGGCGGAGACGAGGTCGCCGCGCTGGCGGCCCTCGCTGAGG 336
```

```
seq_name: gb_est2:BG412166

seq_documentation_block:
LOCUS       BG412166      519 bp    mRNA            EST       13-MAR-2001
DEFINITION  OV2_38_F11.b1_A002 Ovary 2 (OV2) Sorghum bicolor cDNA, mRNA
            sequence.
ACCESSION   BG412166
VERSION     BG412166.1  GI:13317719
KEYWORDS    EST.
SOURCE      sorghum.
  ORGANISM  Sorghum bicolor
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC
            clade; Panicoideae; Andropogoneae; Sorghum.
REFERENCE   1  (bases 1 to 519)
  AUTHORS   Cordonnier-Pratt,M.-M., Gingle,A., Marsala,C., Sudman,M. and Pratt
            ,L.H.
  TITLE     An EST database from Sorghum: ovaries of varying immature stages
  JOURNAL   Unpublished (2000)
COMMENT     Contact: Cordonnier-Pratt MM
            Department of Botany
            The University of Georgia
            Plant Sciences Building, Rm. 2502, Athens, GA 30602-7271, USA
            Tel: 706 542 1860
            Fax: 706 542 1805
            Email: mmpratt@uga.edu
            Seq primer: JEN REV
            High quality sequence stop: 517
            POLYA=No.
FEATURES             Location/Qualifiers
     source          1. .519
                     /organism="Sorghum bicolor"
                     /db_xref="taxon:4558"
                     /clone_lib="Ovary 2 (OV2)"
                     /note="Organ: Mix of ovaries of varying immature stages
                     from 8-week-old plants; Vector: pBluescript II from Lambda
                     Zap II; Site_1: XhoI; Site_2: EcoRI; The library was made
                     from poly-A RNA in the cloning vector lambda ZAP II.
                     Clones to be sequenced were prepared by mass excision."
BASE COUNT       95 a    178 c    162 g     84 t
ORIGIN


alignment_scores:
            Quality:   98.00            Length:     99
              Ratio:  1.815              Gaps:      5
 Percent Similarity:  54.545   Percent Identity: 34.343

alignment_block:
 US-09-508-832-4 x BG412166   ..

 Align seg 1/1  to: BG412166  from: 1  to: 519

       2 AlaLysGlnProSerAspValSerSerGluCysAspArgGlu........ 15
         :::||||||||||         :::||||||||||:::|||
      48 TCGAAACAGCCCCCT......GCGTCCGAATGTTCTAGAGCCTTCCGGCC 91

      16 .......GlyGlyGlnLeuGlnProAlaGlu.ArgProProGlnLeuArg 29
                 |||      :::   |||   ||| |||   |||      |||
      92 GGGCCTAGGGTTCCCGATGGGGCCCGAGGAGGCGGAGGCCGTGCTGGAGA 141

      30 Pro...GlyAlaProThrSerLeuGlnThrGluProGlnAspArgSerPr 45
         |||   |||:::   |||:::   ::::::   |||:::      |||||
     142 CCATCTGGGACCTCCACGACAAGGTCAGCGACGCCATCCACGCCCTCTCC 191

      45 oAlaProMetSerCysAspLysSerThrGlnThrPro.............S 58
         |||||||   ||||||      |||:::   :::|||             |
     192 CGCGCCCACTTCCTGCGCGCCGTCCGCCGCCGCGCCTCCGCCGCGGACAA 241

      58 erProProCysGlnAlaPheAsnHisTyrLeuSerAlaMetAlaSerIle 74
         ||||||||      :::   :::         ::::::   ||||||
     242 GCCCGCCGCCGGCCTCGTGTACGTCAAGGGCGGGGGTCTGGGCCTCGGCG 291
```

```
     75 ArgGlnSerGlnGluGluProGluAspLeuArgProGluIleArg 89
        |||   ::::::   :::            ||||||::::::|||
    292 CGCGCGGCGGAGACGAGGTCGCCGCGCTGGCGGCCCTCGCTGAGG 336

seq_name: gb_est1:BE728552

seq_documentation_block:
LOCUS       BE728552      758 bp    mRNA            EST       15-SEP-2000
DEFINITION  601562025F1 NIH_MGC_20 Homo sapiens cDNA clone IMAGE:3831463 5',
            mRNA sequence.
ACCESSION   BE728552
VERSION     BE728552.1  GI:10142544
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 758)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: ATCC/DCTD/DTP
             cDNA Library Preparation: Ling Hong/Rubin Laboratory
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Incyte Genomics, Inc.
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at: image.llnl.gov
            Plate: LLCM507  row: b  column: 08
            High quality sequence stop: 717.
FEATURES             Location/Qualifiers
     source          1. .758
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:3831463"
                     /clone_lib="NIH_MGC_20"
                     /tissue_type="melanotic melanoma"
                     /lab_host="DH10B (phage-resistant)"
                     /note="Organ: skin; Vector: pOTB7; Site_1: XhoI; Site_2:
                     EcoRI; cDNA made by oligo-dT priming.  Directionally
                     cloned into EcoRI/XhoI sites using the following 5'
                     adaptor: GGCACGAG(G).  Size-selected >500bp for average
                     insert size 1.8kb.  Library constructed by Ling Hong in
                     the laboratory of Gerald M. Rubin (University of
                     California, Berkeley) using  ZAP-cDNA synthesis kit
                     (Stratagene) and Superscript II RT (Life Technologies)."
BASE COUNT      195 a    204 c    232 g    127 t
ORIGIN


alignment_scores:
            Quality:   95.50            Length:    141
              Ratio:  1.404              Gaps:      9
 Percent Similarity:  48.227   Percent Identity: 29.078

alignment_block:
 US-09-508-832-4 x BE728552   ..

 Align seg 1/1  to: BE728552  from: 1  to: 758

     13 AspArgGluGlyGlyGlnLeuGlnProAlaGluArgProProGlnLeuAr 29
        :::||||||||||   :::   :::|||   ||||||||||||   |||
    157 GAACGAGAAGGAGACCATGCAAAGCCT...GAACGACCGCCTGGCCTCTT 203

     29 gProGlyAlaProThrSerLeuGlnThrGluProGlnAspArgSerProA 46
        ||||||              :::   ||||||   ||||||:::|||
    204 ACCTGGACAG............AGTGAGGAGCCTGGAGACCGAGAACCG. 240

     46 laProMetSerCysAspLysSerThrGlnThrPro.............. 57
                 :::   :::   |||:::|||
    241 ..............GAGGCTGGAGAGCAAAATCCGGGAGCACTTGGAGAA 276

     58 ...SerProProCysGlnAlaPheAsnHisTyrLeuSerAlaMetAlaSe 73
           :::.  |||   |||   ::::::||||||::::::   :::
    277 GAAGGGACCCCAGGTCAGAGACTGAGCCATTACTTCAAGATCATC..... 321

     73 rIleArgGlnSerGlnGluGluProGluAspLeuArgProGluIle.... 88
                               ||||||||||||   :::|||
    322 ........................GAGGACCTGAGGGCTCAGATCTTCG 346

     89 ...................ArgIleAlaGlnGluLeu.........Arg 95
                           ||||||:::   ::::::         |||
    347 CAAATACTGTGGACAATGCCCGCATCGTTCTGCAGATTGACAATGCCCGT 396

     96 ArgIleGlyAspGluPheAsnGluThrTyr..............ThrAr 107
           :::|||:::|||:::      |||                  ||
    397 CTTGCTGCTGATGACTTTAGAGTCAAGTATGAGACAGAGCTGGCCATGCG 446

    107 gArgValPheAlaAsnAsp.........TyrArgGluAlaGluAspHisP 121
        |:::      ||||||           |||::::::   |||
    447 CCAGTCTGTGGAGAACGACATCCATGGGCTCCGCAAGGTCATTGATGACA 496

    121 roGlnMetValIleLeuGlnLeu 128
          :::::::::   |||||||||
    497 CCAATATCACACGACTGCAGCTG 519
```

THIS PAGE BLANK (USPTO)

```
OM of: US-09-508-832-6 to: GenEmbl:*   out_format : pfs

Date: Dec 11, 2001  1:45 AM

About: Results were produced by the GenCore software, version 4.5,
       Copyright (c) 1993-2000 Compugen Ltd.

Command line parameters:
 -MODEL=frame+_p2n.model -DEV=xlp
 -Q=/cgn2_1/USPTO_spool/US09508832/runat_10122001_110349_29549/app_query.fasta_1.620
 -DB=GenEmbl -QFMT=fastap -SUFFIX=rge -GAPOP=12.000 -GAPEXT=4.000
 -MINMATCH=0.100 -LOOPCL=0.000 -LOOPEXT=0.000 -QGAPOP=4.500
 -QGAPEXT=0.050 -XGAPOP=10.000 -XGAPEXT=0.500 -FGAPOP=6.000
 -FGAPEXT=7.000 -YGAPOP=10.000 -YGAPEXT=0.500 -DELOP=6.000
 -DELEXT=7.000 -START=1 -MATRIX=blosum62 -TRANS=human40.cdi
 -LIST=45 -DOCALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0
 -ALIGN=15 -MODE=LOCAL -OUTFMT=pfs -NORM=ext -MINLEN=0
 -MAXLEN=2000000000 -USER=US09508832_@CGN1_1_0 -NCPU=6 -ICPU=3
 -LONGLOG -NO_XLPXY -WAIT -THREADS=1

Search information block:
Query: US-09-508-832-6
Query length: 196
Database: GenEmbl:*
Database sequences: 1472140
Database length: -341344837
Search time (sec): 2485.920000


score_list:
Sequence              Strd Orig    ZScore    EScore Len  ! Documentation  ..
gb_pat:AX031283        +  1046.00 1072.64  1.8e-51   590  ! AX031283 Sequence 5 from Patent
gb_un:AX031309         +  1046.00 1072.64  1.8e-51   590  ! AX031309 Sequence 5 from Patent
gb_ro:AF032459         +  1046.00 1072.63  1.8e-51   591  ! AF032459 Mus musculus BimEL mRN
gb_ro:AF065433         +  1031.00 1057.44  1.3e-50   591  ! AF065433 Rattus norvegicus Bcl-
gb_pat:AX031287        +   908.00  932.83  1.1e-43   596  ! AX031287 Sequence 9 from Patent
gb_un:AX031313         +   908.00  932.83  1.1e-43   596  ! AX031313 Sequence 9 from Patent
gb_pr:AF032457         +   908.00  932.82  1.1e-43   597  ! AF032457 Homo sapiens BimEL mRN
gb_pat:AX031281        +   704.00  728.38  2.7e-32   422  ! AX031281 Sequence 3 from Patent
gb_un:AX031307         +   704.00  728.38  2.7e-32   422  ! AX031307 Sequence 3 from Patent
gb_ro:AF032460         +   704.00  728.37  2.7e-32   423  ! AF032460 Mus musculus BimL mRNA
gb_ro:AF136927         +   689.00  713.18  1.9e-31   423  ! AF136927 Rattus norvegicus Bcl-
gb_htg:AC013332        -   656.00  642.39  1.7e-27  180569  ! AC013332 Homo sapiens chromo
gb_pat:AX031285        +   591.00  614.04  6.3e-26   416  ! AX031285 Sequence 7 from Patent
gb_un:AX031311         +   591.00  614.04  6.3e-26   416  ! AX031311 Sequence 7 from Patent
gb_pr:AF032458         +   591.00  614.03  6.3e-26   417  ! AF032458 Homo sapiens BimL mRNA
gb_pat:AX031279        +   521.00  544.55  4.7e-22   332  ! AX031279 Sequence 1 from Patent
gb_un:AX031305         +   521.00  544.55  4.7e-22   332  ! AX031305 Sequence 1 from Patent
gb_ro:AF032461         +   521.00  544.53  4.7e-22   333  ! AF032461 Mus musculus BimS mRNA
gb_ro:AF065432         +   503.00  526.30  4.9e-21   333  ! AF065432 Rattus norvegicus Bcl-
gb_ro:AF065431         +   488.00  512.14  3.0e-20   282  ! AF065431 Rattus norvegicus Bcl-
gb_htg:AC013332        +   137.00  116.83  313.12   180569  ! AC013332 Homo sapiens chromo
gb_htg:AC046192        -   132.00  112.91  517.68   150035  ! AC046192 Homo sapiens chromo
gb_htg:AC046192        +   128.00  108.86  870.33   150035  ! AC046192 Homo sapiens chromo
gb_pat:AX167575        +   123.50  129.68  60.23     2454  ! AX167575 Sequence 75 from Pate
gb_pr:AC010519         +   123.50  111.50  620.37    46740  ! AC010519 Homo sapiens chromos
gb_htg:AC024055        -   123.50  103.40  1.8e+03  173690  ! AC024055 Homo sapiens chromo
gb_htg:AC093324        +   123.50  102.71  1.9e+03  194241  ! AC093324 Homo sapiens chromo
gb_ba:SCBAC1A6         +   122.50  110.64  692.62    45590  ! AL589708 Streptomyces coelico
gb_pr:U82668           -   120.50  104.76  1.5e+03   85132  ! U82668 Homo sapiens shox gene
gb_ba:SC8A6            +   120.00  109.32  820.08    37445  ! AL031013 Streptomyces coelico
gb_ba:SCH10            +   119.00  107.97  974.62    39524  ! AL049754 Streptomyces coelico
gb_htg:AC020922        +   119.00  100.41  2.6e+03  134693  ! AC020922 Homo sapiens chromo
gb_pr:AK027614         +   118.50  124.74  113.47    2405  ! AK027614 Homo sapiens cDNA FLI
gb_ba:STMWHIB12X       +   117.50  119.71  216.42    4615  ! L22864 Streptomyces aureofacie
gb_pr:AC005095         -   117.50   99.60  2.9e+03  119937  ! AC005095 Homo sapiens BAC cl
gb_htg:AC027602        -   117.00   95.43  4.9e+03  217346  ! AC027602 Homo sapiens chromo
gb_htg:AC040922        -   116.50   96.11  4.5e+03  179372  ! AC040922 Homo sapiens chromo
gb_ba:AB044076         -   116.00  117.82  275.56    4896  ! AB044076 Myxococcus xanthus mo
gb_om:RABBCCBIII       +   116.00  115.02  394.85    7713  ! D14157 Rabbit mRNA for rabbit
gb_htg:AC019122        +   116.00   94.70  5.3e+03  207636  ! AC019122 Homo sapiens chromo
gb_pr:AC008403         -   115.50   94.74  5.3e+03  190075  ! AC008403 Homo sapiens chromo
gb_pr:AF283769         +   114.50  121.97  161.92    1955  ! AF283769 Homo sapiens clone
gb_pr:HSM801714        +   114.50  120.89  185.98    2329  ! AL136746 Homo sapiens mRNA;
gb_pr:AF220542         -   114.50  109.23  829.49   15405  ! AF220542 Homo sapiens FcRN
gb_pr:AP002815         +   114.50   94.25  5.7e+03  174484  ! AP002815 Homo sapiens gen

seq_name: gb_pat:AX031283

seq_documentation_block:
LOCUS       AX031283      590 bp    DNA             PAT       20-SEP-2000
DEFINITION  Sequence 5 from Patent WO9914321.
ACCESSION   AX031283
VERSION     AX031283.1  GI:10278614
KEYWORDS    .
SOURCE      unidentified.
  ORGANISM  unidentified
            unclassified.
REFERENCE   1  (bases 1 to 590)
  AUTHORS   O'Reilly,L., Puthalakath,H., Adams,J., O'Connor,L., Cory,S.,
            Huang,D.C. and Strasser,A.
  TITLE     Novel therapeutic molecules
  JOURNAL   Patent: WO 9914321-A 5 25-MAR-1999;
            INST MEDICAL W & E HALL (AU) ; PUTHALAKATH HAMSA (AU) ; REILLY
            LORRAINE O (AU) ; ADAMS JERRY (AU) ; CONNOR LIAM O (AU) ; CORY
            SUZANNE (AU) ; HUANG DAVID C S (AU) ; STRASSER ANDREAS (AU)
FEATURES             Location/Qualifiers
     source          1. .590
                     /organism="unidentified"
                     /db_xref="taxon:32644"
     CDS             1. .>587
                     /note="unnamed protein product"
                     /codon_start=1
                     /protein_id="CAC09655.1"
                     /db_xref="GI:10278615"
                     /translation="MAKQPSDVSSECDREGGQLQPAERPPQLRPGAPTSLQTEPQGNP
                     DGEGDRCPHGSPQGPLAPPASPGPFATRSPLFIFVRRSSLLSRSSSGYFSFDTDRSPA
                     PMSCDKSTQTPSPPCQAFNHYLSAMASIRQSQEEPEDLRPEIRIAQELRRIGDEFNET
                     YTRRVFANDYREAEDHPQMVILQLLRFIFRLVWRR"
BASE COUNT      137 a    178 c    150 g    125 t
ORIGIN


alignment_scores:
           Quality: 1046.00              Length:    196
             Ratio:  5.337                 Gaps:      0
 Percent Similarity: 100.000   Percent Identity: 100.000

alignment_block:
 US-09-508-832-6 x AX031283   ..

 Align seg 1/1  to: AX031283  from: 1  to: 590

       1 MetAlaLysGlnProSerAspValSerSerGluCysAspArgGluGlyGl 17
         ||||||||||||||||||||||||||||||||||||||||||||||||||
       1 ATGGCCAAGCAACCTTCTGATGTAAGTTCTGAGTGTGACAGAGAAGGTGG 50

      17 yGlnLeuGlnProAlaGluArgProProGlnLeuArgProGlyAlaProT 34
         ||||||||||||||||||||||||||||||||||||||||||||||||||
      51 ACAATTGCAGCCTGCTGAGAGGCCTCCCCAGCTCAGGCCTGGGGCCCCTA 100

      34 hrSerLeuGlnThrGluProGlnGlyAsnProAspGlyGluGlyAspArg 50
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     101 CCTCCCTACAGACAGAACCGCAAGGTAATCCCGACGGCGAAGGGGACCGC 150

      51 CysProHisGlySerProGlnGlyProLeuAlaProProAlaSerProGl 67
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     151 TGCCCCCACGGCAGCCCTCAGGGCCCGCTGGCCCCACCGGCCAGCCCTGG 200

      67 yProPheAlaThrArgSerProLeuPheIlePheValArgArgSerSerL 84
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     201 CCCTTTTGCTACCAGATCCCCACTTTTCATCTTTGTGAGAAGATCTTCTC 250

      84 euLeuSerArgSerSerSerGlyTyrPheSerPheAspThrAspArgSer 100
         ||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
     251 TGCTGTCCCGGTCCTCCAGTGGGTATTTCTCTTTTGACACAGACAGGAGC 300

     101 ProAlaProMetSerCysAspLysSerThrGlnThrProSerProProCy 117
         |||||||||||||||||||||||||||||||||||||||||||||||||||
     301 CCGGCACCCATGAGTTGTGACAAGTCAACACAAACCCCAAGTCCTCCTTG 350

     117 sGlnAlaPheAsnHisTyrLeuSerAlaMetAlaSerIleArgGlnSerG 134
         |||||||||||||||||||||||||||||||||||||||||||||||||||
     351 CCAGGCCTTCAACCACTATCTCAGTGCAATGGCTTCCATACGACAGTCTC 400

     134 lnGluGluProGluAspLeuArgProGluIleArgIleAlaGlnGluLeu 150
         |||||||||||||||||||||||||||||||||||||||||||||||||||
     401 AGGAGGAACCTGAAGATCTGCGCCCGGAGATACGGATTGCACAGGAGCTG 450

     151 ArgArgIleGlyAspGluPheAsnGluThrTyrThrArgArgValPheAl 167
         |||||||||||||||||||||||||||||||||||||||||||||||||||
     451 CGGCGGATCGGAGACGAGTTCAACGAAACTTACACAAGGAGGGTGTTTGC 500

     167 aAsnAspTyrArgGluAlaGluAspHisProGlnMetValIleLeuGlnL 184
         |||||||||||||||||||||||||||||||||||||||||||||||||||
     501 AAATGATTACCGCGAGGCTGAAGACCACCCTCAAATGGTTATCTTACAAC 550

     184 euLeuArgPheIlePheArgLeuValTrpArgArgHis 196
         |||||||||||||||||||||||||||||||||||||
     551 TGTTACGCTTTATCTTCCGTCTGGTATGGAGAAGGCAT 588

seq_name: gb_un:AX031309

seq_documentation_block:
LOCUS       AX031309      590 bp    DNA             UNA       20-SEP-2000
DEFINITION  Sequence 5 from Patent WO9914321.
ACCESSION   AX031309
VERSION     AX031309.1  GI:10278637
KEYWORDS    .
SOURCE      unidentified.
  ORGANISM  unidentified
            unclassified.
REFERENCE   1  (bases 1 to 590)
  AUTHORS   O'Reilly,L., Puthalakath,H., Adams,J., O'Connor,L., Cory,S.,
            Huang,D.C. and Strasser,A.
  TITLE     Novel therapeutic molecules
  JOURNAL   Patent: WO 9914321-A 25-MAR-1999;
            INST MEDICAL W & E HALL (AU) ; PUTHALAKATH HAMSA (AU) ; REILLY
            LORRAINE O (AU) ; ADAMS JERRY (AU) ; CONNOR LIAM O (AU) ; CORY
            SUZANNE (AU) ; HUANG DAVID C S (AU) ; STRASSER ANDREAS (AU)
FEATURES             Location/Qualifiers
     source          1. .590
                     /organism="unidentified"
                     /db_xref="taxon:32644"
     CDS             1. .>587
                     /note="unnamed protein product"
                     /codon_start=1
                     /protein_id="CAC09660.1"
                     /db_xref="GI:10278638"
                     /translation="MAKQPSDVSSECDREGGQLQPAERPPQLRPGAPTSLQTEPQGNP
                     DGEGDRCPHGSPQGPLAPPASPGPFATRSPLFIFVRRSSLLSRSSSGYFSFDTDRSPA
                     PMSCDKSTQTPSPPCQAFNHYLSAMASIRQSQEEPEDLRPEIRIAQELRRIGDEFNET
                     YTRRVFANDYREAEDHPQMVILQLLRFIFRLVWRR"
BASE COUNT      137 a    178 c    150 g    125 t
ORIGIN


alignment_scores:
            Quality: 1046.00            Length:    196
              Ratio:  5.337              Gaps:      0
 Percent Similarity: 100.000   Percent Identity: 100.000

alignment_block:
 US-09-508-832-6 x AX031309   ..

 Align seg 1/1  to: AX031309  from: 1  to: 590

       1 MetAlaLysGlnProSerAspValSerSerGluCysAspArgGluGlyGl 17
         |||||||||||||||||||||||||||||||||||||||||||||||||||
       1 ATGGCCAAGCAACCTTCTGATGTAAGTTCTGAGTGTGACAGAGAAGGTGG 50

      17 yGlnLeuGlnProAlaGluArgProProGlnLeuArgProGlyAlaProT 34
         |||||||||||||||||||||||||||||||||||||||||||||||||||
      51 ACAATTGCAGCCTGCTGAGAGGCCTCCCCAGCTCAGGCCTGGGGCCCCTA 100

      34 hrSerLeuGlnThrGluProGlnGlyAsnProAspGlyGluGlyAspArg 50
         |||||||||||||||||||||||||||||||||||||||||||||||||||
     101 CCTCCCTACAGACAGAACCGCAAGGTAATCCCGACGGCGAAGGGGACCGC 150

      51 CysProHisGlySerProGlnGlyProLeuAlaProProAlaSerProGl 67
         |||||||||||||||||||||||||||||||||||||||||||||||||||
     151 TGCCCCCACGGCAGCCCTCAGGGCCCGCTGGCCCCACCGGCCAGCCCTGG 200

      67 yProPheAlaThrArgSerProLeuPheIlePheValArgArgSerSerL 84
         |||||||||||||||||||||||||||||||||||||||||||||||||||
     201 CCCTTTTGCTACCAGATCCCCACTTTTCATCTTTGTGAGAAGATCTTCTC 250

      84 euLeuSerArgSerSerSerGlyTyrPheSerPheAspThrAspArgSer 100
         |||||||||||||||||||||||||||||||||||||||||||||||||||
     251 TGCTGTCCCGGTCCTCCAGTGGGTATTTCTCTTTTGACACAGACAGGAGC 300

     101 ProAlaProMetSerCysAspLysSerThrGlnThrProSerProProCy 117
         |||||||||||||||||||||||||||||||||||||||||||||||||||
     301 CCGGCACCCATGAGTTGTGACAAGTCAACACAAACCCCAAGTCCTCCTTG 350

     117 sGlnAlaPheAsnHisTyrLeuSerAlaMetAlaSerIleArgGlnSerG 134
         |||||||||||||||||||||||||||||||||||||||||||||||||||
     351 CCAGGCCTTCAACCACTATCTCAGTGCAATGGCTTCCATACGACAGTCTC 400

     134 lnGluGluProGluAspLeuArgProGluIleArgIleAlaGlnGluLeu 150
         |||||||||||||||||||||||||||||||||||||||||||||||||||
     401 AGGAGGAACCTGAAGATCTGCGCCCGGAGATACGGATTGCACAGGAGCTG 450

     151 ArgArgIleGlyAspGluPheAsnGluThrTyrThrArgArgValPheAl 167
         |||||||||||||||||||||||||||||||||||||||||||||||||||
     451 CGGCGGATCGGAGACGAGTTCAACGAAACTTACACAAGGAGGGTGTTTGC 500

     167 aAsnAspTyrArgGluAlaGluAspHisProGlnMetValIleLeuGlnL 184
         |||||||||||||||||||||||||||||||||||||||||||||||||||
     501 AAATGATTACCGCGAGGCTGAAGACCACCCTCAAATGGTTATCTTACAAC 550

     184 euLeuArgPheIlePheArgLeuValTrpArgArgHis 196
         |||||||||||||||||||||||||||||||||||||
     551 TGTTACGCTTTATCTTCCGTCTGGTATGGAGAAGGCAT 588

seq_name: gb_ro:AF032459

seq_documentation_block:
LOCUS       AF032459      591 bp    mRNA            ROD       19-FEB-1998
DEFINITION  Mus musculus BimEL mRNA, complete cds.
ACCESSION   AF032459
VERSION     AF032459.1  GI:2895499
KEYWORDS    .
SOURCE      house mouse.
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 591)
  AUTHORS   O'Connor,L., Strasser,A., O'Reilly,L.A., Hausmann,G., Adams,J.M.,
            Cory,S. and Huang,D.C.
  TITLE     Bim: a novel member of the Bcl-2 family that promotes apoptosis
  JOURNAL   EMBO J. 17 (2), 384-395 (1998)
  MEDLINE   98094360
   PUBMED   9430630
REFERENCE   2  (bases 1 to 591)
  AUTHORS   O'Connor,L., Strasser,A., O'Reilly,L.A., Hausmann,G., Adams,J.M.,
            Cory,S. and Huang,D.C.S.
  TITLE     Direct Submission
  JOURNAL   Submitted (03-NOV-1997) Molecular Genetics of Cancer, The Walter &
            Eliza Hall Institute of Medical Research, PO Royal Melbourne
            Hospital, Parkville, Victoria 3050, Australia
```

```
FEATURES             Location/Qualifiers
     source          1. .591
                     /organism="Mus musculus"
                     /db_xref="taxon:10090"
     CDS             1. .591
                     /note="pro-apoptotic BH3-containing Bcl-2 family member"
                     /codon_start=1
                     /product="BimEL"
                     /protein_id="AAC40029.1"
                     /db_xref="GI:2895500"
                     /translation="MAKQPSDVSSECDREGGQLQPAERPPQLRPGAPTSLQTEPQGNP
                     DGEGDRCPHGSPQGPLAPPASPGPFATRSPLFIFVRRSSLLSRSSSGYFSFDTDRSPA
                     PMSCDKSTQTPSPPCQAFNHYLSAMASIRQSQEEPEDLRPEIRIAQELRRIGDEFNET
                     YTRRVFANDYREAEDHPQMVILQLLRFIFRLVWRRH"
BASE COUNT      138 a    178 c    150 g    125 t
ORIGIN


alignment_scores:
           Quality: 1046.00            Length:    196
             Ratio:  5.337               Gaps:      0
 Percent Similarity: 100.000   Percent Identity: 100.000

alignment_block:
 US-09-508-832-6 x AF032459   ..

 Align seg 1/1  to: AF032459  from: 1  to: 591

       1 MetAlaLysGlnProSerAspValSerSerGluCysAspArgGluGlyGl 17
         ||||||||||||||||||||||||||||||||||||||||||||||||||
       1 ATGGCCAAGCAACCTTCTGATGTAAGTTCTGAGTGTGACAGAGAAGGTGG 50

      17 yGlnLeuGlnProAlaGluArgProProGlnLeuArgProGlyAlaProT 34
         ||||||||||||||||||||||||||||||||||||||||||||||||||
      51 ACAATTGCAGCCTGCTGAGAGGCCTCCCCAGCTCAGGCCTGGGGCCCCTA 100

      34 hrSerLeuGlnThrGluProGlnGlyAsnProAspGlyGluGlyAspArg 50
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     101 CCTCCCTACAGACAGAACCGCAAGGTAATCCCGACGGCGAAGGGGACCGC 150

      51 CysProHisGlySerProGlnGlyProLeuAlaProProAlaSerProGl 67
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     151 TGCCCCCACGGCAGCCCTCAGGGCCCGCTGGCCCCACCGGCCAGCCCTGG 200

      67 yProPheAlaThrArgSerProLeuPheIlePheValArgArgSerSerL 84
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     201 CCCTTTTGCTACCAGATCCCCACTTTTCATCTTTGTGAGAAGATCTTCTC 250

      84 euLeuSerArgSerSerSerGlyTyrPheSerPheAspThrAspArgSer 100
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     251 TGCTGTCCCGGTCCTCCAGTGGGTATTTCTCTTTTGACACAGACAGGAGC 300

     101 ProAlaProMetSerCysAspLysSerThrGlnThrProSerProProCy 117
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     301 CCGGCACCCATGAGTTGTGACAAGTCAACACAAACCCCAAGTCCTCCTTG 350

     117 sGlnAlaPheAsnHisTyrLeuSerAlaMetAlaSerIleArgGlnSerG 134
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     351 CCAGGCCTTCAACCACTATCTCAGTGCAATGGCTTCCATACGACAGTCTC 400

     134 lnGluGluProGluAspLeuArgProGluIleArgIleAlaGlnGluLeu 150
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     401 AGGAGGAACCTGAAGATCTGCGCCCGGAGATACGGATTGCACAGGAGCTG 450

     151 ArgArgIleGlyAspGluPheAsnGluThrTyrThrArgArgValPheAl 167
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     451 CGGCGGATCGGAGACGAGTTCAACGAAACTTACACAAGGAGGGTGTTTGC 500

     167 aAsnAspTyrArgGluAlaGluAspHisProGlnMetValIleLeuGlnL 184
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     501 AAATGATTACCGCGAGGCTGAAGACCACCCTCAAATGGTTATCTTACAAC 550

     184 euLeuArgPheIlePheArgLeuValTrpArgArgHis 196
```

```
         ||||||||||||||||||||||||||||||||||||
     551 TGTTACGCTTTATCTTCCGTCTGGTATGGAGAAGGCAT 588

seq_name: gb_ro:AF065433

seq_documentation_block:
LOCUS       AF065433      591 bp    mRNA            ROD       11-MAR-1999
DEFINITION  Rattus norvegicus Bcl-2 related ovarian death gene product BOD-L
            mRNA, complete cds.
ACCESSION   AF065433
VERSION     AF065433.1  GI:3228569
KEYWORDS    .
SOURCE      Norway rat.
  ORGANISM  Rattus norvegicus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
            Rattus.
REFERENCE   1  (bases 1 to 591)
  AUTHORS   Hsu,S.Y., Lin,P. and Hsueh,A.J.
  TITLE     BOD (Bcl-2-related ovarian death gene) is an ovarian BH3
            domain-containing proapoptotic Bcl-2 protein capable of
            dimerization with diverse antiapoptotic Bcl-2 members
  JOURNAL   Mol. Endocrinol. 12 (9), 1432-1440 (1998)
  MEDLINE   98400436
REFERENCE   2  (bases 1 to 591)
  AUTHORS   Hsu,S.Y. and Hsueh,A.J.W.
  TITLE     Direct Submission
  JOURNAL   Submitted (15-MAY-1998) GYN/OB, Stanford University, MSOB S385,
            Stanford, CA 94305, USA
FEATURES             Location/Qualifiers
     source          1. .591
                     /organism="Rattus norvegicus"
                     /db_xref="taxon:10116"
     CDS             1. .591
                     /codon_start=1
                     /product="Bcl-2 related ovarian death gene product BOD-L"
                     /protein_id="AAC23595.1"
                     /db_xref="GI:3228570"
                     /translation="MAKQPSDVNSECDREGGQLQPAERPPQLRPGAPTSLQTESQGNP
                     DGEGDRCPHGSPQGPLAPPASPGPFATRSPLFIFVRRSSLLSRSSSGYFSFDTDRSPA
                     PMSCDKSTQTPSPPCQAFNHYLSAMASIRQSQEEPEDLRPEIRIAQELRRIGDEFNET
                     YTRRAFANDYREAEDHPQMVILQLLRFIFRLVWRRH"
BASE COUNT      139 a    177 c    153 g    122 t
ORIGIN


alignment_scores:
           Quality: 1031.00            Length:    196
             Ratio:  5.287               Gaps:      0
 Percent Similarity:  99.490   Percent Identity: 98.469

alignment_block:
 US-09-508-832-6 x AF065433   ..

 Align seg 1/1  to: AF065433  from: 1  to: 591

       1 MetAlaLysGlnProSerAspValSerSerGluCysAspArgGluGlyGl 17
         ||||||||||||||||||||||:::||||||||||||||||||||||||||
       1 ATGGCCAAGCAACCTTCTGATGTAAATTCTGAGTGTGACAGAGAAGGTGG 50

      17 yGlnLeuGlnProAlaGluArgProProGlnLeuArgProGlyAlaProT 34
         ||||||||||||||||||||||||||||||||||||||||||||||||||
      51 ACAATTGCAGCCTGCTGAGAGGCCTCCCCAGCTCAGGCCTGGGGCCCCTA 100

      34 hrSerLeuGlnThrGluProGlnGlyAsnProAspGlyGluGlyAspArg 50
         ||||||||||||||||  ||||||||||||||||||||||||||||||||
     101 CCTCCCTACAGACAGAATCGCAAGGTAATCCCGACGGCGAAGGGGACCGC 150

      51 CysProHisGlySerProGlnGlyProLeuAlaProProAlaSerProGl 67
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     151 TGCCCCCACGGCAGCCCTCAGGGCCCGCTGGCCCCACCGGCCAGCCCTGG 200

      67 yProPheAlaThrArgSerProLeuPheIlePheValArgArgSerSerL 84
```

```
     ||||||||||||||||||||||||||||||||||||||||||||||||||
201 TCCTTTTGCTACCAGATCCCCACTTTTCATCTTTGTGAGAAGATCTTCTC 250

 84 euLeuSerArgSerSerSerGlyTyrPheSerPheAspThrAspArgSer 100
    ||||||||||||||||||||||||||||||||||||||||||||||||||
251 TGCTGTCCCGGTCCTCCAGTGGGTATTTCTCTTTTGACACAGACAGGAGC 300

101 ProAlaProMetSerCysAspLysSerThrGlnThrProSerProProCy 117
    ||||||||||||||||||||||||||||||||||||||||||||||||||
301 CCGGCACCCATGAGTTGTGACAAGTCAACACAAACCCCAAGTCCTCCTTG 350

117 sGlnAlaPheAsnHisTyrLeuSerAlaMetAlaSerIleArgGlnSerG 134
    ||||||||||||||||||||||||||||||||||||||||||||||||||
351 CCAGGCCTTCAACCATTATCTCAGTGCAATGGCTTCCATAAGGCAGTCTC 400

134 lnGluGluProGluAspLeuArgProGluIleArgIleAlaGlnGluLeu 150
    ||||||||||||||||||||||||||||||||||||||||||||||||||
401 AGGAGGAACCTGAAGATCTGCGCCCAGAGATACGGATCGCACAGGAGCTG 450

151 ArgArgIleGlyAspGluPheAsnGluThrTyrThrArgArgValPheAl 167
    ||||||||||||||||||||||||||||||||||||||||||:::|||||
451 CGGCGGATCGGAGACGAGTTCAATGAGACTTACACGAGGAGGGCGTTTGC 500

167 aAsnAspTyrArgGluAlaGluAspHisProGlnMetValIleLeuGlnL 184
    ||||||||||||||||||||||||||||||||||||||||||||||||||
501 AAACGATTACCGAGAGGCGGAAGACCACCCGCAAATGGTTATCTTACAAC 550

184 euLeuArgPheIlePheArgLeuValTrpArgArgHis 196
    |||||||||||||||||||||||||||||||||||
551 TGTTACGATTCATCTTCCGTCTGGTCTGGAGAAGGCAC 588

seq_name: gb_pat:AX031287

seq_documentation_block:
LOCUS       AX031287      596 bp    DNA             PAT       20-SEP-2000
DEFINITION  Sequence 9 from Patent WO9914321.
ACCESSION   AX031287
VERSION     AX031287.1  GI:10278618
KEYWORDS    .
SOURCE      unidentified.
  ORGANISM  unidentified
            unclassified.
REFERENCE   1  (bases 1 to 596)
  AUTHORS   O'Reilly,L., Puthalakath,H., Adams,J., O'Connor,L., Cory,S.,
            Huang,D.C. and Strasser,A.
  TITLE     Novel therapeutic molecules
  JOURNAL   Patent: WO 9914321-A 9 25-MAR-1999;
            INST MEDICAL W & E HALL (AU) ; PUTHALAKATH HAMSA (AU) ; REILLY
            LORRAINE O (AU) ; ADAMS JERRY (AU) ; CONNOR LIAM O (AU) ; CORY
            SUZANNE (AU) ; HUANG DAVID C S (AU) ; STRASSER ANDREAS (AU)
FEATURES             Location/Qualifiers
     source          1. .596
                     /organism="unidentified"
                     /db_xref="taxon:32644"
     CDS             1. .>594
                     /note="unnamed protein product"
                     /codon_start=1
                     /protein_id="CAC09657.1"
                     /db_xref="GI:10278619"
                     /translation="MAKQPSDVSSECDREGRQLQPAERPPQLRPGAPTSLQTEPQGNP
                     EGNHGGEGDSCPHGSPQGPLAPPASPGPFATRSPLFIFMRRSSLLSRSSSGYFSFDTD
                     RSPAPMSCDKSTQTPSPPCQAFNHYLSAMASMRQAEPADMRPEIWIAQELRRIGDEFN
                     AYYARRVFLNNYQAAEDHPRMVILRLLRYIVRLVWRMH"
BASE COUNT      145 a    175 c    146 g    130 t
ORIGIN


alignment_scores:
            Quality:  908.00            Length:    200
              Ratio:  4.935               Gaps:      2
 Percent Similarity:  92.000   Percent Identity: 86.500

alignment_block:
```

```
US-09-508-832-6 x AX031287  ..

Align seg 1/1  to: AX031287  from: 1  to: 596

      1 MetAlaLysGlnProSerAspValSerSerGluCysAspArgGluGlyGl 17
        ||||||||||||||||||||||||||||||||||||||||||||||||||
      1 ATGGCAAAGCAACCTTCTGATGTAAGTTCTGAGTGTGACCGAGAAGGTAG 50

     17 yGlnLeuGlnProAlaGluArgProProGlnLeuArgProGlyAlaProT 34
        ||||||||||||||||||||||||||||||||||||||||||||||||||
     51 ACAATTGCAGCCTGCGGAGAGGCCTCCCCAGCTCAGACCTGGGGCCCCTA 100

     34 hrSerLeuGlnThrGluProGlnGlyAsnProAsp.............Gly 46
        ||||||||||||||||||||||||||||||||:::              |||
    101 CCTCCCTACAGACAGAGCCACAAGGTAATCCTGAAGGCAATCACGGAGGT 150

     47 GluGlyAspArgCysProHisGlySerProGlnGlyProLeuAlaProPr 63
        |||||||||   ||||||||||||||||||||||||||||||||||||||
    151 GAAGGGGACAGCTGCCCCCACGGCAGCCCTCAGGGCCCGCTGGCCCCACC 200

     63 oAlaSerProGlyProPheAlaThrArgSerProLeuPheIlePheValA 80
        ||||||||||||||||||||||||||||||||||||||||||||||:::|
    201 TGCCAGCCCTGGCCCTTTTGCTACCAGATCCCCGCTTTTCATCTTTATGA 250

     80 rgArgSerSerLeuLeuSerArgSerSerSerGlyTyrPheSerPheAsp 96
        ||||||||||||||||||||||||||||||||||||||||||||||||||
    251 GAAGATCCTCCCTGCTGTCTCGATCCTCCAGTGGGTATTTCTCTTTTGAC 300

     97 ThrAspArgSerProAlaProMetSerCysAspLysSerThrGlnThrPr 113
        ||||||||||||||||||||||||||||||||||||||||||||||||||
    301 ACAGACAGGAGCCCAGCACCCATGAGTTGTGACAAATCAACACAAACCCC 350

    113 oSerProProCysGlnAlaPheAsnHisTyrLeuSerAlaMetAlaSerI 130
        ||||||||||||||||||||||||||||||||||||||||||||||||:
    351 AAGTCCTCCTTGCCAGGCCTTCAACCACTATCTCAGTGCAATGGCTTCCA 400

    130 leArgGlnSerGlnGluGluProGluAspLeuArgProGluIleArgIle 146
        ::||||||:::      ||||||   |||:::||||||||||||   |||
    401 TGAGGCAGGCT......GAACCTGCAGATATGCGCCCAGAGATATGGATC 444

    147 AlaGlnGluLeuArgArgIleGlyAspGluPheAsnGluThrTyrThrAr 163
        ||||||||||||||||||||||||||||||||      |||:::||
    445 GCCCAAGAGTTGCGGCGTATCGGAGACGAGTTTAACGCTTACTATGCAAG 494

    163 gArgValPheAlaAsnAspTyrArgGluAlaGluAspHisProGlnMetV 180
        ||||||||||   |||:::|||:::   ||||||||||||||:::||||
    495 GAGGGTATTTTTGAATAATTACCAAGCAGCCGAAGACCACCCACGAATGG 544

    180 alIleLeuGlnLeuLeuArgPheIlePheArgLeuValTrpArgArgHis 196
        ||||||||:::|||||||||:::|||   |||||||||||||||   |||
    545 TTATCTTACGACTGTTACGTTACATTGTCCGCCTGGTGTGGAGAATGCAT 594

seq_name: gb_un:AX031313

seq_documentation_block:
LOCUS       AX031313      596 bp    DNA             UNA       20-SEP-2000
DEFINITION  Sequence 9 from Patent WO9914321.
ACCESSION   AX031313
VERSION     AX031313.1  GI:10278641
KEYWORDS    .
SOURCE      unidentified.
  ORGANISM  unidentified
            unclassified.
REFERENCE   1  (bases 1 to 596)
  AUTHORS   O'Reilly,L., Puthalakath,H., Adams,J., O'Connor,L., Cory,S.,
            Huang,D.C. and Strasser,A.
  TITLE     Novel therapeutic molecules
  JOURNAL   Patent: WO 9914321-A 25-MAR-1999;
            INST MEDICAL W & E HALL (AU) ; PUTHALAKATH HAMSA (AU) ; REILLY
            LORRAINE O (AU) ; ADAMS JERRY (AU) ; CONNOR LIAM O (AU) ; CORY
            SUZANNE (AU) ; HUANG DAVID C S (AU) ; STRASSER ANDREAS (AU)
FEATURES             Location/Qualifiers
     source          1. .596
```

```
                     /organism="unidentified"
                     /db_xref="taxon:32644"
     CDS             1. .>594
                     /note="unnamed protein product"
                     /codon_start=1
                     /protein_id="CAC09662.1"
                     /db_xref="GI:10278642"
                     /translation="MAKQPSDVSSECDREGRQLQPAERPPQLRPGAPTSLQTEPQGNP
                     EGNHGGEGDSCPHGSPQGPLAPPASPGPFATRSPLFIFMRRSSLLSRSSSGYFSFDTD
                     RSPAPMSCDKSTQTPSPPCQAFNHYLSAMASMRQAEPADMRPEIWIAQELRRIGDEFN
                     AYYARRVFLNNYQAAEDHPRMVILRLLRYIVRLVWRMH"
BASE COUNT      145 a    175 c    146 g    130 t
ORIGIN


alignment_scores:
            Quality:  908.00             Length:    200
              Ratio:  4.935                Gaps:      2
 Percent Similarity:  92.000   Percent Identity: 86.500

alignment_block:
 US-09-508-832-6 x AX031313   ..

 Align seg 1/1  to: AX031313  from: 1  to: 596

      1 MetAlaLysGlnProSerAspValSerSerGluCysAspArgGluGlyGl 17
        ||||||||||||||||||||||||||||||||||||||||||||||||||
      1 ATGGCAAAGCAACCTTCTGATGTAAGTTCTGAGTGTGACCGAGAAGGTAG 50

     17 yGlnLeuGlnProAlaGluArgProProGlnLeuArgProGlyAlaProT 34
        ||||||||||||||||||||||||||||||||||||||||||||||||||
     51 ACAATTGCAGCCTGCGGAGAGGCCTCCCCAGCTCAGACCTGGGGCCCCTA 100

     34 hrSerLeuGlnThrGluProGlnGlyAsnProAsp.....Gly 46
        ||||||||||||||||||||||||||||||||:::         |||
    101 CCTCCCTACAGACAGAGCCACAAGGTAATCCTGAAGGCAATCACGGAGGT 150

     47 GluGlyAspArgCysProHisGlySerProGlnGlyProLeuAlaProPr 63
        |||||||||   ||||||||||||||||||||||||||||||||||||||
    151 GAAGGGGACAGCTGCCCCCACGGCAGCCCTCAGGGCCCGCTGGCCCCACC 200

     63 oAlaSerProGlyProPheAlaThrArgSerProLeuPheIlePheValA 80
        ||||||||||||||||||||||||||||||||||||||||||||||||:::|
    201 TGCCAGCCCTGGCCCTTTTGCTACCAGATCCCCGCTTTTCATCTTTATGA 250

     80 rgArgSerSerLeuLeuSerArgSerSerSerGlyTyrPheSerPheAsp 96
        ||||||||||||||||||||||||||||||||||||||||||||||||||
    251 GAAGATCCTCCCTGCTGTCTCGATCCTCCAGTGGGTATTTCTCTTTTGAC 300

     97 ThrAspArgSerProAlaProMetSerCysAspLysSerThrGlnThrPr 113
        ||||||||||||||||||||||||||||||||||||||||||||||||||
    301 ACAGACAGGAGCCCAGCACCCATGAGTTGTGACAAATCAACACAAACCCC 350

    113 oSerProProCysGlnAlaPheAsnHisTyrLeuSerAlaMetAlaSerI 130
        ||||||||||||||||||||||||||||||||||||||||||||||||||:
    351 AAGTCCTCCTTGCCAGGCCTTCAACCACTATCTCAGTGCAATGGCTTCCA 400

    130 leArgGlnSerGlnGluGluProGluAspLeuArgProGluIleArgIle 146
        ::|||||:::      ||||||   |||:::|||||||||||||   |||
    401 TGAGGCAGGCT......GAACCTGCAGATATGCGCCCAGAGATATGGATC 444

    147 AlaGlnGluLeuArgArgIleGlyAspGluPheAsnGluThrTyrThrAr 163
        |||||||||||||||||||||||||||||||||      |||:::||
    445 GCCCAAGAGTTGCGGCGTATCGGAGACGAGTTTAACGCTTACTATGCAAG 494

    163 gArgValPheAlaAsnAspTyrArgGluAlaGluAspHisProGlnMetV 180
        ||||||||||   |||:::|||:::   |||||||||||||:::||||
    495 GAGGGTATTTTTGAATAATTACCAAGCAGCCGAAGACCACCCACGAATGG 544

    180 alIleLeuGlnLeuLeuArgPheIlePheArgLeuValTrpArgArgHis 196
        ||||||||:::|||||||||:::|||   ||||||||||||||   |||
    545 TTATCTTACGACTGTTACGTTACATTGTCCGCCTGGTGTGGAGAATGCAT 594
```

```
seq_name: gb_pr:AF032457

seq_documentation_block:
LOCUS       AF032457      597 bp    mRNA            PRI       19-FEB-1998
DEFINITION  Homo sapiens BimEL mRNA, complete cds.
ACCESSION   AF032457
VERSION     AF032457.1  GI:2895495
KEYWORDS    .
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 597)
  AUTHORS   O'Connor,L., Strasser,A., O'Reilly,L.A., Hausmann,G., Adams,J.M.,
            Cory,S. and Huang,D.C.
  TITLE     Bim: a novel member of the Bcl-2 family that promotes apoptosis
  JOURNAL   EMBO J. 17 (2), 384-395 (1998)
  MEDLINE   98094360
   PUBMED   9430630
REFERENCE   2  (bases 1 to 597)
  AUTHORS   O'Connor,L., Strasser,A., O'Reilly,L.A., Hausmann,G., Adams,J.M.,
            Cory,S. and Huang,D.C.S.
  TITLE     Direct Submission
  JOURNAL   Submitted (03-NOV-1997) Molecular Genetics of Cancer, The Walter &
            Eliza Hall Institute of Medical Research, PO Royal Melbourne
            Hospital, Parkville, Victoria 3050, Australia
FEATURES             Location/Qualifiers
     source          1. .597
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
     CDS             1. .597
                     /note="pro-apoptotic BH3-containing Bcl-2 family member"
                     /codon_start=1
                     /product="BimEL"
                     /protein_id="AAC39593.1"
                     /db_xref="GI:2895496"
                     /translation="MAKQPSDVSSECDREGRQLQPAERPPQLRPGAPTSLQTEPQGNP
                     EGNHGGEGDSCPHGSPQGPLAPPASPGPFATRSPLFIFMRRSSLLSRSSSGYFSFDTD
                     RSPAPMSCDKSTQTPSPPCQAFNHYLSAMASMRQAEPADMRPEIWIAQELRRIGDEFN
                     AYYARRVFLNNYQAAEDHPRMVILRLLRYIVRLVWRMH"
BASE COUNT      146 a    175 c    146 g    130 t
ORIGIN


alignment_scores:
            Quality:  908.00             Length:    200
              Ratio:  4.935                Gaps:      2
 Percent Similarity:  92.000   Percent Identity: 86.500

alignment_block:
 US-09-508-832-6 x AF032457   ..

 Align seg 1/1  to: AF032457  from: 1  to: 597

      1 MetAlaLysGlnProSerAspValSerSerGluCysAspArgGluGlyGl 17
        ||||||||||||||||||||||||||||||||||||||||||||||||||
      1 ATGGCAAAGCAACCTTCTGATGTAAGTTCTGAGTGTGACCGAGAAGGTAG 50

     17 yGlnLeuGlnProAlaGluArgProProGlnLeuArgProGlyAlaProT 34
        ||||||||||||||||||||||||||||||||||||||||||||||||||
     51 ACAATTGCAGCCTGCGGAGAGGCCTCCCCAGCTCAGACCTGGGGCCCCTA 100

     34 hrSerLeuGlnThrGluProGlnGlyAsnProAsp.....Gly 46
        ||||||||||||||||||||||||||||||||:::         |||
    101 CCTCCCTACAGACAGAGCCACAAGGTAATCCTGAAGGCAATCACGGAGGT 150

     47 GluGlyAspArgCysProHisGlySerProGlnGlyProLeuAlaProPr 63
        |||||||||   ||||||||||||||||||||||||||||||||||||||
    151 GAAGGGGACAGCTGCCCCCACGGCAGCCCTCAGGGCCCGCTGGCCCCACC 200

     63 oAlaSerProGlyProPheAlaThrArgSerProLeuPheIlePheValA 80
        ||||||||||||||||||||||||||||||||||||||||||||||||:::|
    201 TGCCAGCCCTGGCCCTTTTGCTACCAGATCCCCGCTTTTCATCTTTATGA 250
```

```
     80 rgArgSerSerLeuLeuSerArgSerSerSerGlyTyrPheSerPheAsp 96
        ||||||||||||||||||||||||||||||||||||||||||||||||||
    251 GAAGATCCTCCCTGCTGTCTCGATCCTCCAGTGGGTATTTCTCTTTTGAC 300

     97 ThrAspArgSerProAlaProMetSerCysAspLysSerThrGlnThrPr 113
        ||||||||||||||||||||||||||||||||||||||||||||||||||
    301 ACAGACAGGAGCCCAGCACCCATGAGTTGTGACAAATCAACACAAACCCC 350

    113 oSerProProCysGlnAlaPheAsnHisTyrLeuSerAlaMetAlaSerI 130
        ||||||||||||||||||||||||||||||||||||||||||||||||||:
    351 AAGTCCTCCTTGCCAGGCCTTCAACCACTATCTCAGTGCAATGGCTTCCA 400

    130 leArgGlnSerGlnGluGluProGluAspLeuArgProGluIleArgIle 146
        ::||||||:::      ||||||   |||:::||||||||||||   |||
    401 TGAGGCAGGCT......GAACCTGCAGATATGCGCCCAGAGATATGGATC 444

    147 AlaGlnGluLeuArgArgIleGlyAspGluPheAsnGluThrTyrThrAr 163
        ||||||||||||||||||||||||||||||||||      |||:::||
    445 GCCCAAGAGTTGCGGCGTATCGGAGACGAGTTTAACGCTTACTATGCAAG 494

    163 gArgValPheAlaAsnAspTyrArgGluAlaGluAspHisProGlnMetV 180
        ||||||||||   |||:::|||:::   ||||||||||||||:::||||
    495 GAGGGTATTTTTGAATAATTACCAAGCAGCCGAAGACCACCCACGAATGG 544

    180 alIleLeuGlnLeuLeuArgPheIlePheArgLeuValTrpArgArgHis 196
        ||||||||:::|||||||||:::|||   ||||||||||||||   |||
    545 TTATCTTACGACTGTTACGTTACATTGTCCGCCTGGTGTGGAGAATGCAT 594

seq_name: gb_pat:AX031281

seq_documentation_block:
LOCUS       AX031281      422 bp    DNA             PAT       20-SEP-2000
DEFINITION  Sequence 3 from Patent WO9914321.
ACCESSION   AX031281
VERSION     AX031281.1  GI:10278612
KEYWORDS    .
SOURCE      unidentified.
  ORGANISM  unidentified
            unclassified.
REFERENCE   1  (bases 1 to 422)
  AUTHORS   O'Reilly,L., Puthalakath,H., Adams,J., O'Connor,L., Cory,S.,
            Huang,D.C. and Strasser,A.
  TITLE     Novel therapeutic molecules
  JOURNAL   Patent: WO 9914321-A 3 25-MAR-1999;
            INST MEDICAL W & E HALL (AU) ; PUTHALAKATH HAMSA (AU) ; REILLY
            LORRAINE O (AU) ; ADAMS JERRY (AU) ; CONNOR LIAM O (AU) ; CORY
            SUZANNE (AU) ; HUANG DAVID C S (AU) ; STRASSER ANDREAS (AU)
FEATURES             Location/Qualifiers
     source          1. .422
                     /organism="unidentified"
                     /db_xref="taxon:32644"
     CDS             1. .>420
                     /note="unnamed protein product"
                     /codon_start=1
                     /protein_id="CAC09654.1"
                     /db_xref="GI:10278613"
                     /translation="MAKQPSDVSSECDREGGQLQPAERPPQLRPGAPTSLQTEPQDRS
                     PAPMSCDKSTQTPSPPCQAFNHYLSAMASIRQSQEEPEDLRPEIRIAQELRRIGDEFN
                     ETYTRRVFANDYREAEDHPQMVILQLLRFIFRLVWRRH"
BASE COUNT      112 a    116 c    109 g     85 t
ORIGIN


alignment_scores:
            Quality:  704.00             Length:    196
              Ratio:  5.029                Gaps:      1
 Percent Similarity:  71.429   Percent Identity: 71.429

alignment_block:
 US-09-508-832-6 x AX031281   ..

 Align seg 1/1  to: AX031281  from: 1  to: 422
```

```
      1 MetAlaLysGlnProSerAspValSerSerGluCysAspArgGluGlyGl 17
        ||||||||||||||||||||||||||||||||||||||||||||||||||
      1 ATGGCCAAGCAACCTTCTGATGTAAGTTCTGAGTGTGACAGAGAAGGTGG 50

     17 yGlnLeuGlnProAlaGluArgProProGlnLeuArgProGlyAlaProT 34
        ||||||||||||||||||||||||||||||||||||||||||||||||||
     51 ACAATTGCAGCCTGCTGAGAGGCCTCCCCAGCTCAGGCCTGGGGCCCCTA 100

     34 hrSerLeuGlnThrGluProGlnGlyAsnProAspGlyGluGlyAspArg 50
        ||||||||||||||||||||||
    101 CCTCCCTACAGACAGAACCGCAA..................................... 123

     51 CysProHisGlySerProGlnGlyProLeuAlaProProAlaSerProGl 67

    123 ..................................................... 123

     67 yProPheAlaThrArgSerProLeuPheIlePheValArgArgSerSerL 84

    123 ..................................................... 123

     84 euLeuSerArgSerSerSerGlyTyrPheSerPheAspThrAspArgSer 100
                                                 |||||||||
    124 ..................................GACAGGAGC 132

    101 ProAlaProMetSerCysAspLysSerThrGlnThrProSerProProCy 117
        ||||||||||||||||||||||||||||||||||||||||||||||||||
    133 CCGGCACCCATGAGTTGTGACAAGTCAACACAAACCCCAAGTCCTCCTTG 182

    117 sGlnAlaPheAsnHisTyrLeuSerAlaMetAlaSerIleArgGlnSerG 134
        ||||||||||||||||||||||||||||||||||||||||||||||||||
    183 CCAGGCCTTCAACCACTATCTCAGTGCAATGGCTTCCATACGACAGTCTC 232

    134 lnGluGluProGluAspLeuArgProGluIleArgIleAlaGlnGluLeu 150
        ||||||||||||||||||||||||||||||||||||||||||||||||||
    233 AGGAGGAACCTGAAGATCTGCGCCCGGAGATACGGATTGCACAGGAGCTG 282

    151 ArgArgIleGlyAspGluPheAsnGluThrTyrThrArgArgValPheAl 167
        ||||||||||||||||||||||||||||||||||||||||||||||||||
    283 CGGCGGATCGGAGACGAGTTCAACGAAACTTACACAAGGAGGGTGTTTGC 332

    167 aAsnAspTyrArgGluAlaGluAspHisProGlnMetValIleLeuGlnL 184
        ||||||||||||||||||||||||||||||||||||||||||||||||||
    333 AAATGATTACCGCGAGGCTGAAGACCACCCTCAAATGGTTATCTTACAAC 382

    184 euLeuArgPheIlePheArgLeuValTrpArgArgHis 196
        ||||||||||||||||||||||||||||||||||||
    383 TGTTACGCTTTATCTTCCGTCTGGTATGGAGAAGGCAT 420

seq_name: gb_un:AX031307

seq_documentation_block:
LOCUS       AX031307      422 bp    DNA             UNA       20-SEP-2000
DEFINITION  Sequence 3 from Patent WO9914321.
ACCESSION   AX031307
VERSION     AX031307.1  GI:10278635
KEYWORDS    .
SOURCE      unidentified.
  ORGANISM  unidentified
            unclassified.
REFERENCE   1  (bases 1 to 422)
  AUTHORS   O'Reilly,L., Puthalakath,H., Adams,J., O'Connor,L., Cory,S.,
            Huang,D.C. and Strasser,A.
  TITLE     Novel therapeutic molecules
  JOURNAL   Patent: WO 9914321-A 25-MAR-1999;
            INST MEDICAL W & E HALL (AU) ; PUTHALAKATH HAMSA (AU) ; REILLY
            LORRAINE O (AU) ; ADAMS JERRY (AU) ; CONNOR LIAM O (AU) ; CORY
            SUZANNE (AU) ; HUANG DAVID C S (AU) ; STRASSER ANDREAS (AU)
FEATURES             Location/Qualifiers
     source          1. .422
                     /organism="unidentified"
                     /db_xref="taxon:32644"
     CDS             1. .>420
```

```
                     /note="unnamed protein product"
                     /codon_start=1
                     /protein_id="CAC09659.1"
                     /db_xref="GI:10278636"
                     /translation="MAKQPSDVSSECDREGGQLQPAERPPQLRPGAPTSLQTEPQDRS
                     PAPMSCDKSTQTPSPPCQAFNHYLSAMASIRQSQEEPEDLRPEIRIAQELRRIGDEFN
                     ETYTRRVFANDYREAEDHPQMVILQLLRFIFRLVWRRH"
BASE COUNT      112 a    116 c    109 g     85 t
ORIGIN


alignment_scores:
            Quality:  704.00             Length:    196
              Ratio:  5.029                Gaps:      1
 Percent Similarity:  71.429   Percent Identity: 71.429

alignment_block:
 US-09-508-832-6 x AX031307   ..

 Align seg 1/1  to: AX031307  from: 1  to: 422

       1 MetAlaLysGlnProSerAspValSerSerGluCysAspArgGluGlyGl 17
         ||||||||||||||||||||||||||||||||||||||||||||||||||
       1 ATGGCCAAGCAACCTTCTGATGTAAGTTCTGAGTGTGACAGAGAAGGTGG 50

      17 yGlnLeuGlnProAlaGluArgProProGlnLeuArgProGlyAlaProT 34
         ||||||||||||||||||||||||||||||||||||||||||||||||||
      51 ACAATTGCAGCCTGCTGAGAGGCCTCCCCAGCTCAGGCCTGGGGCCCCTA 100

      34 hrSerLeuGlnThrGluProGlnGlyAsnProAspGlyGluGlyAspArg 50
         ||||||||||||||||||||
     101 CCTCCCTACAGACAGAACCGCAA..................................... 123

      51 CysProHisGlySerProGlnGlyProLeuAlaProProAlaSerProGl 67

     123 ..................................................... 123

      67 yProPheAlaThrArgSerProLeuPheIlePheValArgArgSerSerL 84

     123 ..................................................... 123

      84 euLeuSerArgSerSerSerGlyTyrPheSerPheAspThrAspArgSer 100
                                                 |||||||||
     124 ..........................................GACAGGAGC 132

     101 ProAlaProMetSerCysAspLysSerThrGlnThrProSerProProCy 117
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     133 CCGGCACCCATGAGTTGTGACAAGTCAACACAAACCCCAAGTCCTCCTTG 182

     117 sGlnAlaPheAsnHisTyrLeuSerAlaMetAlaSerIleArgGlnSerG 134
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     183 CCAGGCCTTCAACCACTATCTCAGTGCAATGGCTTCCATACGACAGTCTC 232

     134 lnGluGluProGluAspLeuArgProGluIleArgIleAlaGlnGluLeu 150
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     233 AGGAGGAACCTGAAGATCTGCGCCCGGAGATACGGATTGCACAGGAGCTG 282

     151 ArgArgIleGlyAspGluPheAsnGluThrTyrThrArgArgValPheAl 167
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     283 CGGCGGATCGGAGACGAGTTCAACGAAACTTACACAAGGAGGGTGTTTGC 332

     167 aAsnAspTyrArgGluAlaGluAspHisProGlnMetValIleLeuGlnL 184
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     333 AAATGATTACCGCGAGGCTGAAGACCACCCTCAAATGGTTATCTTACAAC 382

     184 euLeuArgPheIlePheArgLeuValTrpArgArgHis 196
         ||||||||||||||||||||||||||||||||||
     383 TGTTACGCTTTATCTTCCGTCTGGTATGGAGAAGGCAT 420

seq_name: gb_ro:AF032460

seq_documentation_block:
LOCUS       AF032460      423 bp    mRNA            ROD       19-FEB-1998
DEFINITION  Mus musculus BimL mRNA, complete cds.
ACCESSION   AF032460
VERSION     AF032460.1  GI:2895501
KEYWORDS    .
SOURCE      house mouse.
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 423)
  AUTHORS   O'Connor,L., Strasser,A., O'Reilly,L.A., Hausmann,G., Adams,J.M.,
            Cory,S. and Huang,D.C.
  TITLE     Bim: a novel member of the Bcl-2 family that promotes apoptosis
  JOURNAL   EMBO J. 17 (2), 384-395 (1998)
  MEDLINE   98094360
   PUBMED   9430630
REFERENCE   2  (bases 1 to 423)
  AUTHORS   O'Connor,L., Strasser,A., O'Reilly,L.A., Hausmann,G., Adams,J.M.,
            Cory,S. and Huang,D.C.S.
  TITLE     Direct Submission
  JOURNAL   Submitted (03-NOV-1997) Molecular Genetics of Cancer, The Walter &
            Eliza Hall Institute of Medical Research, PO Royal Melbourne
            Hospital, Parkville, Victoria 3050, Australia
FEATURES             Location/Qualifiers
     source          1. .423
                     /organism="Mus musculus"
                     /db_xref="taxon:10090"
     CDS             1. .423
                     /note="pro-apoptotic BH3-containing Bcl-2 family member"
                     /codon_start=1
                     /product="BimL"
                     /protein_id="AAC40030.1"
                     /db_xref="GI:2895502"
                     /translation="MAKQPSDVSSECDREGGQLQPAERPPQLRPGAPTSLQTEPQDRS
                     PAPMSCDKSTQTPSPPCQAFNHYLSAMASIRQSQEEPEDLRPEIRIAQELRRIGDEFN
                     ETYTRRVFANDYREAEDHPQMVILQLLRFIFRLVWRRH"
BASE COUNT      113 a    116 c    109 g     85 t
ORIGIN


alignment_scores:
            Quality:  704.00             Length:    196
              Ratio:  5.029                Gaps:      1
 Percent Similarity:  71.429   Percent Identity: 71.429

alignment_block:
 US-09-508-832-6 x AF032460   ..

 Align seg 1/1  to: AF032460  from: 1  to: 423

       1 MetAlaLysGlnProSerAspValSerSerGluCysAspArgGluGlyGl 17
         ||||||||||||||||||||||||||||||||||||||||||||||||||
       1 ATGGCCAAGCAACCTTCTGATGTAAGTTCTGAGTGTGACAGAGAAGGTGG 50

      17 yGlnLeuGlnProAlaGluArgProProGlnLeuArgProGlyAlaProT 34
         ||||||||||||||||||||||||||||||||||||||||||||||||||
      51 ACAATTGCAGCCTGCTGAGAGGCCTCCCCAGCTCAGGCCTGGGGCCCCTA 100

      34 hrSerLeuGlnThrGluProGlnGlyAsnProAspGlyGluGlyAspArg 50
         ||||||||||||||||||||
     101 CCTCCCTACAGACAGAACCGCAA..................................... 123

      51 CysProHisGlySerProGlnGlyProLeuAlaProProAlaSerProGl 67

     123 ..................................................... 123

      67 yProPheAlaThrArgSerProLeuPheIlePheValArgArgSerSerL 84

     123 ..................................................... 123

      84 euLeuSerArgSerSerSerGlyTyrPheSerPheAspThrAspArgSer 100
                                                 |||||||||
     124 ..........................................GACAGGAGC 132
```

```
     101 ProAlaProMetSerCysAspLysSerThrGlnThrProSerProProCy 117
         ||||||||||||||||||||||||||||||||||||||||||||||||||||
     133 CCGGCACCCATGAGTTGTGACAAGTCAACACAAACCCCAAGTCCTCCTTG 182

     117 sGlnAlaPheAsnHisTyrLeuSerAlaMetAlaSerIleArgGlnSerG 134
         ||||||||||||||||||||||||||||||||||||||||||||||||||||
     183 CCAGGCCTTCAACCACTATCTCAGTGCAATGGCTTCCATACGACAGTCTC 232

     134 lnGluGluProGluAspLeuArgProGluIleArgIleAlaGlnGluLeu 150
         ||||||||||||||||||||||||||||||||||||||||||||||||||||
     233 AGGAGGAACCTGAAGATCTGCGCCCGGAGATACGGATTGCACAGGAGCTG 282

     151 ArgArgIleGlyAspGluPheAsnGluThrTyrThrArgArgValPheAl 167
         ||||||||||||||||||||||||||||||||||||||||||||||||||||
     283 CGGCGGATCGGAGACGAGTTCAACGAAACTTACACAAGGAGGGTGTTTGC 332

     167 aAsnAspTyrArgGluAlaGluAspHisProGlnMetValIleLeuGlnL 184
         ||||||||||||||||||||||||||||||||||||||||||||||||||||
     333 AAATGATTACCGCGAGGCTGAAGACCACCCTCAAATGGTTATCTTACAAC 382

     184 euLeuArgPheIlePheArgLeuValTrpArgArgHis 196
         ||||||||||||||||||||||||||||||||||||
     383 TGTTACGCTTTATCTTCCGTCTGGTATGGAGAAGGCAT 420

seq_name: gb_ro:AF136927

seq_documentation_block:
LOCUS       AF136927      423 bp    mRNA            ROD       21-APR-1999
DEFINITION  Rattus norvegicus Bcl-2 related apoptotic gene product BimL (bimL)
            mRNA, complete cds.
ACCESSION   AF136927
VERSION     AF136927.1  GI:4590514
KEYWORDS    .
SOURCE      Norway rat.
  ORGANISM  Rattus norvegicus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
            Rattus.
REFERENCE   1  (bases 1 to 423)
  AUTHORS   Chen,D., Simon,R.P. and Chen,J.
  TITLE     Cloning of rat bimEL and bimL, and their differential expression in
            ischemia and normal rat brain
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 423)
  AUTHORS   Chen,D., Simon,R.P. and Chen,J.
  TITLE     Direct Submission
  JOURNAL   Submitted (24-MAR-1999) Department of Neurology, BST, S-526,
            Pittsburgh University Medical School, 3500 Terrace Street,
            Pittsburgh, PA 15213, USA
FEATURES             Location/Qualifiers
     source          1. .423
                     /organism="Rattus norvegicus"
                     /db_xref="taxon:10116"
     gene            1. .423
                     /gene="bimL"
     CDS             1. .423
                     /gene="bimL"
                     /codon_start=1
                     /product="Bcl-2 related apoptotic gene product BimL"
                     /protein_id="AAD26594.1"
                     /db_xref="GI:4590515"
                     /translation="MAKQPSDVNSECDREGGQLQPAERPPQLRPGAPTSLQTESQDRS
                     PAPMSCDKSTQTPSPPCQAFNHYLSAMASIRQSQEEPEDLRPEIRIAQELRRIGDEFN
                     ETYTRRAFANDYREAEDHPQMVILQLLRFIFRLVWRRH"
BASE COUNT      116 a    113 c    112 g     82 t
ORIGIN


alignment_scores:
           Quality:  689.00            Length:    196
             Ratio:  4.957               Gaps:      1
 Percent Similarity:  70.918   Percent Identity: 69.898
```

```
alignment_block:
 US-09-508-832-6 x AF136927   ..

 Align seg 1/1  to: AF136927  from: 1  to: 423

       1 MetAlaLysGlnProSerAspValSerSerGluCysAspArgGluGlyGl 17
         ||||||||||||||||||||||||:::||||||||||||||||||||||||
       1 ATGGCCAAGCAACCTTCTGATGTAAATTCTGAGTGTGACAGAGAAGGTGG 50

      17 yGlnLeuGlnProAlaGluArgProProGlnLeuArgProGlyAlaProT 34
         ||||||||||||||||||||||||||||||||||||||||||||||||||||
      51 ACAATTGCAGCCTGCTGAGAGGCCTCCCCAGCTCAGGCCTGGGGCCCCTA 100

      34 hrSerLeuGlnThrGluProGlnGlyAsnProAspGlyGluGlyAspArg 50
         ||||||||||||||||||   |||
     101 CCTCCCTACAGACAGAATCGCAA..................................... 123

      51 CysProHisGlySerProGlnGlyProLeuAlaProProAlaSerProGl 67

     123 ..................................................... 123

      67 yProPheAlaThrArgSerProLeuPheIlePheValArgArgSerSerL 84

     123 ..................................................... 123

      84 euLeuSerArgSerSerSerGlyTyrPheSerPheAspThrAspArgSer 100
                                                  |||||||||
     124 .....................................GACAGGAGC 132

     101 ProAlaProMetSerCysAspLysSerThrGlnThrProSerProProCy 117
         ||||||||||||||||||||||||||||||||||||||||||||||||||||
     133 CCGGCACCCATGAGTTGTGACAAGTCAACACAAACCCCAAGTCCTCCTTG 182

     117 sGlnAlaPheAsnHisTyrLeuSerAlaMetAlaSerIleArgGlnSerG 134
         ||||||||||||||||||||||||||||||||||||||||||||||||||||
     183 CCAGGCCTTCAACCATTATCTTAGCGCAATGGCTTCCATAAGGCAGTCTC 232

     134 lnGluGluProGluAspLeuArgProGluIleArgIleAlaGlnGluLeu 150
         ||||||||||||||||||||||||||||||||||||||||||||||||||||
     233 AGGAGGAACCTGAAGATCTGCGCCCAGAGATACGGATCGCACAGGAGCTG 282

     151 ArgArgIleGlyAspGluPheAsnGluThrTyrThrArgArgValPheAl 167
         ||||||||||||||||||||||||||||||||||||||||||||:::|||||
     283 CGGAGGATCGGAGACGAGTTCAATGAGACTTACACGAGGAGGGCGTTTGC 332

     167 aAsnAspTyrArgGluAlaGluAspHisProGlnMetValIleLeuGlnL 184
         ||||||||||||||||||||||||||||||||||||||||||||||||||||
     333 AAACGATTACCGAGAGGCGGAAGACCACCCGCAAATGGTTATCTTACAAC 382

     184 euLeuArgPheIlePheArgLeuValTrpArgArgHis 196
         ||||||||||||||||||||||||||||||||||||
     383 TGTTACGATTCATCTTCCGTCTGGTATGGAGAAGGCAT 420

seq_name: gb_htg:AC013332

seq_documentation_block:
LOCUS       AC013332   180569 bp    DNA             HTG       16-MAR-2000
DEFINITION  Homo sapiens chromosome 2 clone RP11-438K19 map 2, WORKING DRAFT
            SEQUENCE, 6 unordered pieces.
ACCESSION   AC013332
VERSION     AC013332.4  GI:7248987
KEYWORDS    HTG; HTGS_PHASE1; HTGS_DRAFT.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 180569)
  AUTHORS   Birren,B., Linton,L., Nusbaum,C. and Lander,E.
  TITLE     Homo sapiens chromosome 2, clone RP11-438K19
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 180569)
  AUTHORS   Birren,B., Linton,L., Nusbaum,C., Lander,E., Allen,N., Anderson,M.,
            Baldwin,J., Barna,N., Beckerly,R., Boguslavkiy,L., Boukhgalter,B.,
```

```
            Brown,A., Castle,A., Colangelo,M., Collins,S., Collymore,A.,
            Cooke,P., DeArellano,K., Dewar,K., Domino,M., Donelan,L., Doyle,M.,
            Ferreira,P., FitzHugh,W., Forrest,C., Funke,R., Gage,D.,
            Galagan,J., Gardyna,S., Grant,G., Hagos,B., Heaford,A., Horton,L.,
            Howland,J.C., Johnson,R., Jones,C., Kann,L., Karatas,A., Klein,J.,
            Lehoczky,J., Lieu,C., Locke,K., Macdonald,P., Marquis,N.,
            McEwan,P., McGurk,A., McKernan,K., McLaughlin,J., Meldrim,J.,
            Morrow,J., Naylor,J., Norman,C.H., O'Connor,T., O'Donnell,P.,
            Peterson,K., Pollara,V., Riley,R., Roy,A., Santos,R., Severy,P.,
            Stange-Thomann,N., Stojanovic,N., Subramanian,A., Talamas,J.,
            Tesfaye,S., Tirrell,A., Vassiliev,H., Vo,A., Wheeler,J., Wu,X.,
            Wyman,D., Ye,W.J., Zimmer,A. and Zody,M.
  TITLE     Direct Submission
  JOURNAL   Submitted (06-NOV-1999) Whitehead Institute/MIT Center for Genome
            Research, 320 Charles Street, Cambridge, MA 02141, USA
COMMENT     On Mar 16, 2000 this sequence version replaced gi:6478985.
            All repeats were identified using RepeatMasker:
            Smit, A.F.A. & Green, P. (1996-1997)
            http://ftp.genome.washington.edu/RM/RepeatMasker.html
            -------------- Genome Center
                Center: Whitehead Institute/ MIT Center for Genome Research
                Center code: WIBR
                Web site: http://www-seq.wi.mit.edu
                Contact: sequence_submissions@genome.wi.mit.edu
            -------------- Project Information
                Center project name: L3811
                Center clone name: 438_K_19
            -------------- Summary Statistics
                Sequencing vector: M13; M77815; 100% of reads
                Chemistry: Dye-terminator Big Dye; 100% of reads
                Assembly program: Phrap; version 0.960731
                Consensus quality: 158049 bases at least Q40
                Consensus quality: 172735 bases at least Q30
                Consensus quality: 178014 bases at least Q20
                Insert size: 175000; agarose-fp
                Insert size: 180069; sum-of-contigs
                Quality coverage: 4.7 in Q20 bases; agarose-fp
                Quality coverage: 4.6 in Q20 bases; sum-of-contigs
            --------------
            * NOTE: This is a 'working draft' sequence. It currently
            * consists of 6 contigs. The true order of the pieces
            * is not known and their order in this sequence record is
            * arbitrary. Gaps between the contigs are represented as
            * runs of N, but the exact sizes of the gaps are unknown.
            * This record will be updated with the finished sequence
            * as soon as it is available and the accession number will
            * be preserved.
            *       1     2165: contig of 2165 bp in length
            *    2166 2265: gap of      100 bp
            *    2266    12624: contig of 10359 bp in length
            *   12625 12724: gap of      100 bp
            *   12725    27927: contig of 15203 bp in length
            *   27928 28027: gap of      100 bp
            *   28028    65158: contig of 37131 bp in length
            *   65159 65258: gap of      100 bp
            *   65259   107505: contig of 42247 bp in length
            *  107506 107605: gap of      100 bp
            *  107606   180569: contig of 72964 bp in length.
FEATURES            Location/Qualifiers
     source         1. .180569
                    /organism="Homo sapiens"
                    /db_xref="taxon:9606"
                    /chromosome="2"
                    /map="2"
                    /clone="RP11-438K19"
                    /clone_lib="RPCI-11 Human Male BAC"
     misc_feature   1. .2165
                    /note="assembly_fragment"
     misc_feature   2266. .12624
                    /note="assembly_fragment
                    clone_end:SP6
                    vector_side:right"
     misc_feature   12725. .27927
                    /note="assembly_fragment"
     misc_feature   28028. .65158
                    /note="assembly_fragment"
     misc_feature   65259. .107505
                    /note="assembly_fragment"
     misc_feature   107606. .180569
                    /note="assembly_fragment
                    clone_end:T7
                    vector_side:right"
BASE COUNT    49889 a  39169 c  39952 g  51056 t    503 others
ORIGIN


alignment_scores:
          Quality:  656.00            Length:    131
            Ratio:  5.248               Gaps:      1
 Percent Similarity:  95.420   Percent Identity: 93.893

alignment_block:
 US-09-508-832-6 x AC013332/rev   ..

 Align seg 1/1  to reverse of: AC013332  from: 1  to: 180569

      1 MetAlaLysGlnProSerAspValSerSerGluCysAspArgGluGlyGl 17
        ||||||||||||||||||||||||||||||||||||||||||||||||||
  15100 ATGGCAAAGCAACCTTCTGATGTAAGTTCTGAGTGTGACCGAGAAGGTAG 15051

     17 yGlnLeuGlnProAlaGluArgProProGlnLeuArgProGlyAlaProT 34
        ||||||||||||||||||||||||||||||||||||||||||||||||||
  15050 ACAATTGCAGCCTGCGGAGAGGCCTCCCCAGCTCAGACCTGGGGCCCCTA 15001

     34 hrSerLeuGlnThrGluProGlnGlyAsnProAsp.....Gly 46
        ||||||||||||||||||||||||||||||:::          |||
  15000 CCTCCCTACAGACAGAGCCACAAGGTAATCCTGAAGGCAATCACGGAGGT 14951

     47 GluGlyAspArgCysProHisGlySerProGlnGlyProLeuAlaProPr 63
        |||||||||  ||||||||||||||||||||||||||||||||||||||||
  14950 GAAGGGGACAGCTGCCCCCACGGCAGCCCTCAGGGCCCGCTGGCCCCACC 14901

     63 oAlaSerProGlyProPheAlaThrArgSerProLeuPheIlePheValA 80
        ||||||||||||||||||||||||||||||||||||||||||||||:::|
  14900 TGCCAGCCCTGGCCCTTTTGCTACCAGATCCCCGCTTTTCATCTTTATGA 14851

     80 rgArgSerSerLeuLeuSerArgSerSerSerGlyTyrPheSerPheAsp 96
        ||||||||||||||||||||||||||||||||||||||||||||||||||
  14850 GAAGATCCTCCCTGCTGTCTCGATCCTCCAGTGGGTATTTCTCTTTTGAC 14801

     97 ThrAspArgSerProAlaProMetSerCysAspLysSerThrGlnThrPr 113
        ||||||||||||||||||||||||||||||||||||||||||||||||||
  14800 ACAGACAGGAGCCCAGCACCCATGAGTTGTGACAAATCAACACAAACCCC 14751

    113 oSerProProCysGlnAlaPheAsnHisTyrLeuSerAlaMet 127
        |||||||||||||||||||||||||||||||||||||||||||
  14750 AAGTCCTCCTTGCCAGGCCTTCAACCACTATCTCAGTGCAATG 14708

seq_name: gb_pat:AX031285

seq_documentation_block:
LOCUS       AX031285      416 bp    DNA             PAT       20-SEP-2000
DEFINITION  Sequence 7 from Patent WO9914321.
ACCESSION   AX031285
VERSION     AX031285.1  GI:10278616
KEYWORDS    .
SOURCE      unidentified.
  ORGANISM  unidentified
            unclassified.
REFERENCE   1  (bases 1 to 416)
  AUTHORS   O'Reilly,L., Puthalakath,H., Adams,J., O'Connor,L., Cory,S.,
            Huang,D.C. and Strasser,A.
  TITLE     Novel therapeutic molecules
  JOURNAL   Patent: WO 9914321-A 7 25-MAR-1999;
            INST MEDICAL W & E HALL (AU) ; PUTHALAKATH HAMSA (AU) ; REILLY
            LORRAINE O (AU) ; ADAMS JERRY (AU) ; CONNOR LIAM O (AU) ; CORY
```

```
           SUZANNE (AU) ; HUANG DAVID C S (AU) ; STRASSER ANDREAS (AU)
FEATURES             Location/Qualifiers
     source          1. .416
                     /organism="unidentified"
                     /db_xref="taxon:32644"
     CDS             1. .>414
                     /note="unnamed protein product"
                     /codon_start=1
                     /protein_id="CAC09656.1"
                     /db_xref="GI:10278617"
                     /translation="MAKQPSDVSSECDREGRQLQPAERPPQLRPGAPTSLQTEPQDRS
                     PAPMSCDKSTQTPSPPCQAFNHYLSAMASMRQAEPADMRPEIWIAQELRRIGDEFNAY
                     YARRVFLNNYQAAEDHPRMVILRLLRYIVRLVWRMH"
BASE COUNT      113 a    113 c    103 g     87 t
ORIGIN


alignment_scores:
           Quality:  591.00            Length:    196
             Ratio:  4.581               Gaps:      2
 Percent Similarity:  65.816   Percent Identity: 61.224

alignment_block:
 US-09-508-832-6 x AX031285   ..

 Align seg 1/1  to: AX031285  from: 1  to: 416

       1 MetAlaLysGlnProSerAspValSerSerGluCysAspArgGluGlyGl 17
         ||||||||||||||||||||||||||||||||||||||||||||||||||
       1 ATGGCAAAGCAACCTTCTGATGTAAGTTCTGAGTGTGACCGAGAAGGTAG 50

      17 yGlnLeuGlnProAlaGluArgProProGlnLeuArgProGlyAlaProT 34
         ||||||||||||||||||||||||||||||||||||||||||||||||||
      51 ACAATTGCAGCCTGCGGAGAGGCCTCCCCAGCTCAGACCTGGGGCCCCTA 100

      34 hrSerLeuGlnThrGluProGlnGlyAsnProAspGlyGluGlyAspArg 50
         ||||||||||||||||||||||
     101 CCTCCCTACAGACAGAGCCACAA..... 123

      51 CysProHisGlySerProGlnGlyProLeuAlaProProAlaSerProGl 67

     123 ..................................................... 123

      67 yProPheAlaThrArgSerProLeuPheIlePheValArgArgSerSerL 84

     123 ..................................................... 123

      84 euLeuSerArgSerSerSerGlyTyrPheSerPheAspThrAspArgSer 100
                                                 |||||||||
     124 .....................................GACAGGAGC 132

     101 ProAlaProMetSerCysAspLysSerThrGlnThrProSerProProCy 117
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     133 CCAGCACCCATGAGTTGTGACAAATCAACACAAACCCCAAGTCCTCCTTG 182

     117 sGlnAlaPheAsnHisTyrLeuSerAlaMetAlaSerIleArgGlnSerG 134
         ||||||||||||||||||||||||||||||||||:::||||||:::
     183 CCAGGCCTTCAACCACTATCTCAGTGCAATGGCTTCCATGAGGCAGGCT. 231

     134 lnGluGluProGluAspLeuArgProGluIleArgIleAlaGlnGluLeu 150
              ||||||   |||:::|||||||||||   ||||||||||||||||
     232 .....GAACCTGCAGATATGCGCCCAGAGATATGGATCGCCCAAGAGTTG 276

     151 ArgArgIleGlyAspGluPheAsnGluThrTyrThrArgArgValPheAl 167
         ||||||||||||||||||||||      |||:::|||||||||||||
     277 CGGCGTATCGGAGACGAGTTTAACGCTTACTATGCAAGGAGGGTATTTTT 326

     167 aAsnAspTyrArgGluAlaGluAspHisProGlnMetValIleLeuGlnL 184
          |||:::|||:::   ||||||||||||||:::||||||||||||:::|
     327 GAATAATTACCAAGCAGCCGAAGACCACCCACGAATGGTTATCTTACGAC 376

     184 euLeuArgPheIlePheArgLeuValTrpArgArgHis 196
         ||||||||:::|||   ||||||||||||||   |||
```

```
     377 TGTTACGTTACATTGTCCGCCTGGTGTGGAGAATGCAT 414

seq_name: gb_un:AX031311

seq_documentation_block:
LOCUS       AX031311      416 bp    DNA             UNA       20-SEP-2000
DEFINITION  Sequence 7 from Patent WO9914321.
ACCESSION   AX031311
VERSION     AX031311.1  GI:10278639
KEYWORDS    .
SOURCE      unidentified.
  ORGANISM  unidentified
            unclassified.
REFERENCE   1  (bases 1 to 416)
  AUTHORS   O'Reilly,L., Puthalakath,H., Adams,J., O'Connor,L., Cory,S.,
            Huang,D.C. and Strasser,A.
  TITLE     Novel therapeutic molecules
  JOURNAL   Patent: WO 9914321-A 25-MAR-1999;
            INST MEDICAL W & E HALL (AU) ; PUTHALAKATH HAMSA (AU) ; REILLY
            LORRAINE O (AU) ; ADAMS JERRY (AU) ; CONNOR LIAM O (AU) ; CORY
            SUZANNE (AU) ; HUANG DAVID C S (AU) ; STRASSER ANDREAS (AU)
FEATURES             Location/Qualifiers
     source          1. .416
                     /organism="unidentified"
                     /db_xref="taxon:32644"
     CDS             1. .>414
                     /note="unnamed protein product"
                     /codon_start=1
                     /protein_id="CAC09661.1"
                     /db_xref="GI:10278640"
                     /translation="MAKQPSDVSSECDREGRQLQPAERPPQLRPGAPTSLQTEPQDRS
                     PAPMSCDKSTQTPSPPCQAFNHYLSAMASMRQAEPADMRPEIWIAQELRRIGDEFNAY
                     YARRVFLNNYQAAEDHPRMVILRLLRYIVRLVWRMH"
BASE COUNT      113 a    113 c    103 g     87 t
ORIGIN


alignment_scores:
           Quality:  591.00            Length:    196
             Ratio:  4.581               Gaps:      2
 Percent Similarity:  65.816   Percent Identity: 61.224

alignment_block:
 US-09-508-832-6 x AX031311   ..

 Align seg 1/1  to: AX031311  from: 1  to: 416

       1 MetAlaLysGlnProSerAspValSerSerGluCysAspArgGluGlyGl 17
         ||||||||||||||||||||||||||||||||||||||||||||||||||
       1 ATGGCAAAGCAACCTTCTGATGTAAGTTCTGAGTGTGACCGAGAAGGTAG 50

      17 yGlnLeuGlnProAlaGluArgProProGlnLeuArgProGlyAlaProT 34
         ||||||||||||||||||||||||||||||||||||||||||||||||||
      51 ACAATTGCAGCCTGCGGAGAGGCCTCCCCAGCTCAGACCTGGGGCCCCTA 100

      34 hrSerLeuGlnThrGluProGlnGlyAsnProAspGlyGluGlyAspArg 50
         ||||||||||||||||||||||
     101 CCTCCCTACAGACAGAGCCACAA..... 123

      51 CysProHisGlySerProGlnGlyProLeuAlaProProAlaSerProGl 67

     123 ..................................................... 123

      67 yProPheAlaThrArgSerProLeuPheIlePheValArgArgSerSerL 84

     123 ..................................................... 123

      84 euLeuSerArgSerSerSerGlyTyrPheSerPheAspThrAspArgSer 100
                                                 |||||||||
     124 .....................................GACAGGAGC 132

     101 ProAlaProMetSerCysAspLysSerThrGlnThrProSerProProCy 117
         ||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
     133 CCAGCACCCATGAGTTGTGACAAATCAACACAAACCCCAAGTCCTCCTTG 182

     117 sGlnAlaPheAsnHisTyrLeuSerAlaMetAlaSerIleArgGlnSerG 134
         ||||||||||||||||||||||||||||||||||||:::|||||||:::
     183 CCAGGCCTTCAACCACTATCTCAGTGCAATGGCTTCCATGAGGCAGGCT. 231

     134 lnGluGluProGluAspLeuArgProGluIleArgIleAlaGlnGluLeu 150
              ||||||  |||:::|||||||||||||   ||||||||||||||||
     232 .....GAACCTGCAGATATGCGCCCAGAGATATGGATCGCCCAAGAGTTG 276

     151 ArgArgIleGlyAspGluPheAsnGluThrTyrThrArgArgValPheAl 167
         ||||||||||||||||||||||||      |||:::|||||||||||||
     277 CGGCGTATCGGAGACGAGTTTAACGCTTACTATGCAAGGAGGGTATTTTT 326

     167 aAsnAspTyrArgGluAlaGluAspHisProGlnMetValIleLeuGlnL 184
          |||:::|||:::   ||||||||||||||||:::|||||||||||:::|
     327 GAATAATTACCAAGCAGCCGAAGACCACCCACGAATGGTTATCTTACGAC 376

     184 euLeuArgPheIlePheArgLeuValTrpArgArgHis 196
         ||||||||:::|||   ||||||||||||||   |||
     377 TGTTACGTTACATTGTCCGCCTGGTGTGGAGAATGCAT 414

seq_name: gb_pr:AF032458

seq_documentation_block:
LOCUS       AF032458      417 bp    mRNA            PRI       19-FEB-1998
DEFINITION  Homo sapiens BimL mRNA, complete cds.
ACCESSION   AF032458
VERSION     AF032458.1  GI:2895497
KEYWORDS    .
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 417)
  AUTHORS   O'Connor,L., Strasser,A., O'Reilly,L.A., Hausmann,G., Adams,J.M.,
            Cory,S. and Huang,D.C.
  TITLE     Bim: a novel member of the Bcl-2 family that promotes apoptosis
  JOURNAL   EMBO J. 17 (2), 384-395 (1998)
  MEDLINE   98094360
   PUBMED   9430630
REFERENCE   2  (bases 1 to 417)
  AUTHORS   O'Connor,L., Strasser,A., O'Reilly,L.A., Hausmann,G., Adams,J.M.,
            Cory,S. and Huang,D.C.S.
  TITLE     Direct Submission
  JOURNAL   Submitted (03-NOV-1997) Molecular Genetics of Cancer, The Walter &
            Eliza Hall Institute of Medical Research, PO Royal Melbourne
            Hospital, Parkville, Victoria 3050, Australia
FEATURES             Location/Qualifiers
     source          1. .417
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
     CDS             1. .417
                     /note="pro-apoptotic BH3-containing Bcl-2 family member"
                     /codon_start=1
                     /product="BimL"
                     /protein_id="AAC39594.1"
                     /db_xref="GI:2895498"
                     /translation="MAKQPSDVSSECDREGRQLQPAERPPQLRPGAPTSLQTEPQDRS
                     PAPMSCDKSTQTPSPPCQAFNHYLSAMASMRQAEPADMRPEIWIAQELRRIGDEFNAY
                     YARRVFLNNYQAAEDHPRMVILRLLRYIVRLVWRMH"
BASE COUNT      114 a    113 c    103 g     87 t
ORIGIN


alignment_scores:
            Quality:  591.00            Length:    196
              Ratio:  4.581               Gaps:      2
 Percent Similarity:  65.816   Percent Identity: 61.224

alignment_block:
 US-09-508-832-6 x AF032458  ..
```

```
Align seg 1/1  to: AF032458  from: 1  to: 417

      1 MetAlaLysGlnProSerAspValSerSerGluCysAspArgGluGlyGl 17
        ||||||||||||||||||||||||||||||||||||||||||||||||||
      1 ATGGCAAAGCAACCTTCTGATGTAAGTTCTGAGTGTGACCGAGAAGGTAG 50

     17 yGlnLeuGlnProAlaGluArgProProGlnLeuArgProGlyAlaProT 34
        ||||||||||||||||||||||||||||||||||||||||||||||||||
     51 ACAATTGCAGCCTGCGGAGAGGCCTCCCCAGCTCAGACCTGGGGCCCCTA 100

     34 hrSerLeuGlnThrGluProGlnGlyAsnProAspGlyGluGlyAspArg 50
        ||||||||||||||||||||
    101 CCTCCCTACAGACAGAGCCACAA..... 123

     51 CysProHisGlySerProGlnGlyProLeuAlaProProAlaSerProGl 67

    123 ..... 123

     67 yProPheAlaThrArgSerProLeuPheIlePheValArgArgSerSerL 84

    123 ..... 123

     84 euLeuSerArgSerSerSerGlyTyrPheSerPheAspThrAspArgSer 100
                                                 |||||||||
    124 .....GACAGGAGC 132

    101 ProAlaProMetSerCysAspLysSerThrGlnThrProSerProProCy 117
        ||||||||||||||||||||||||||||||||||||||||||||||||||
    133 CCAGCACCCATGAGTTGTGACAAATCAACACAAACCCCAAGTCCTCCTTG 182

    117 sGlnAlaPheAsnHisTyrLeuSerAlaMetAlaSerIleArgGlnSerG 134
        ||||||||||||||||||||||||||||||||||||:::|||||||:::
    183 CCAGGCCTTCAACCACTATCTCAGTGCAATGGCTTCCATGAGGCAGGCT. 231

    134 lnGluGluProGluAspLeuArgProGluIleArgIleAlaGlnGluLeu 150
             ||||||  |||:::|||||||||||||   ||||||||||||||||
    232 .....GAACCTGCAGATATGCGCCCAGAGATATGGATCGCCCAAGAGTTG 276

    151 ArgArgIleGlyAspGluPheAsnGluThrTyrThrArgArgValPheAl 167
        ||||||||||||||||||||||||      |||:::|||||||||||||
    277 CGGCGTATCGGAGACGAGTTTAACGCTTACTATGCAAGGAGGGTATTTTT 326

    167 aAsnAspTyrArgGluAlaGluAspHisProGlnMetValIleLeuGlnL 184
         |||:::|||:::   ||||||||||||||||:::|||||||||||:::|
    327 GAATAATTACCAAGCAGCCGAAGACCACCCACGAATGGTTATCTTACGAC 376

    184 euLeuArgPheIlePheArgLeuValTrpArgArgHis 196
        ||||||||:::|||   ||||||||||||||   |||
    377 TGTTACGTTACATTGTCCGCCTGGTGTGGAGAATGCAT 414
```

OM of: US-09-508-832-6 to: N_Geneseq_1101:* out_format : pfs

Date: Dec 11, 2001 1:50 AM

About: Results were produced by the GenCore software, version 4.5,
Copyright (c) 1993-2000 Compugen Ltd.

Command line parameters:
-MODEL=frame+_p2n.model -DEV=xlp
-Q=/cgn2_1/USPTO_spool/US09508832/runat_10122001_110349_29569/app_query.fasta_1.620
-DB=N_Geneseq_1101 -QFMT=fastap -SUFFIX=rng -GAPOP=12.000
-GAPEXT=4.000 -MINMATCH=0.100 -LOOPCL=0.000 -LOOPEXT=0.000
-QGAPOP=4.500 -QGAPEXT=0.050 -XGAPOP=10.000 -XGAPEXT=0.500
-FGAPOP=6.000 -FGAPEXT=7.000 -YGAPOP=10.000 -YGAPEXT=0.500
-DELOP=6.000 -DELEXT=7.000 -START=1 -MATRIX=blosum62
-TRANS=human40.cdi -LIST=45 -DOCALIGN=200 -THR_SCORE=pct
-THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=pfs
-NORM=ext -MINLEN=0 -MAXLEN=2000000000
-USER=US09508832_@CGN1_1_327 -NCPU=6 -ICPU=3 -LONGLOG -NO_XLPXY
-WAIT -THREADS=1

Search information block:
Query: US-09-508-832-6
Query length: 196
Database: N_Geneseq_1101:*
Database sequences: 930621
Database length: 428662619
Search time (sec): 249.590000

score_list:

Sequence	Strd	Orig	ZScore	EScore	Len	! Documentation ..
/SIDS2/gcgdata/geneseq/geneseqn/NA1999.DAT:AAX24995	+	1046.00	1245.75	2.6e-61	590	!
/SIDS2/gcgdata/geneseq/geneseqn/NA1999.DAT:AAX24997	+	908.00	1082.07	3.4e-52	596	!
/SIDS2/gcgdata/geneseq/geneseqn/NA1999.DAT:AAX24994	+	704.00	843.04	7.0e-39	422	!
/SIDS2/gcgdata/geneseq/geneseqn/NA1999.DAT:AAX24996	+	591.00	709.19	2.0e-31	416	!
/SIDS2/gcgdata/geneseq/geneseqn/NA1999.DAT:AAX24993	+	521.00	628.05	6.6e-27	332	!
/SIDS2/gcgdata/geneseq/geneseqn/NA1998.DAT:AAV35620	-	120.50	115.96	220.71	32367	!
/SIDS2/gcgdata/geneseq/geneseqn/NA2001.DAT:AAH17812	+	118.50	134.76	19.80	2405	!
/SIDS2/gcgdata/geneseq/geneseqn/NA2000.DAT:AAZ61771	+	112.50	131.93	28.46	1421	!
/SIDS2/gcgdata/geneseq/geneseqn/NA2001.DAT:AAC99704	+	112.50	131.93	28.46	1421	!
/SIDS2/gcgdata/geneseq/geneseqn/NA2001.DAT:AAD07131	-	110.00	123.44	84.53	2800	!
/SIDS2/gcgdata/geneseq/geneseqn/NA2001.DAT:AAD07132	-	110.00	123.44	84.62	2803	!
/SIDS2/gcgdata/geneseq/geneseqn/NA2001.DAT:AAD07130	-	110.00	123.43	84.69	2805	!
/SIDS2/gcgdata/geneseq/geneseqn/NA1996.DAT:AAT18023	+	109.50	116.42	208.19	6171	!
/SIDS2/gcgdata/geneseq/geneseqn/NA2000.DAT:AAA60103	+	109.50	116.42	208.19	6171	!
/SIDS2/gcgdata/geneseq/geneseqn/NA2000.DAT:AAZ50085	+	109.00	128.83	42.35	1249	!
/SIDS2/gcgdata/geneseq/geneseqn/NA2001.DAT:AAI36331	-	108.50	124.44	74.37	1991	!
/SIDS2/gcgdata/geneseq/geneseqn/NA2001.DAT:AAI49460	-	108.50	121.56	107.65	2837	!
/SIDS2/gcgdata/geneseq/geneseqn/NA1996.DAT:AAT18020	+	108.50	115.64	229.96	5868	!
/SIDS2/gcgdata/geneseq/geneseqn/NA2000.DAT:AAA60104	+	108.50	115.64	229.96	5868	!
/SIDS2/gcgdata/geneseq/geneseqn/NA2000.DAT:AAA60123	+	108.50	115.64	229.96	5868	!
/SIDS2/gcgdata/geneseq/geneseqn/NA2000.DAT:AAA60124	+	108.50	115.64	229.96	5868	!
/SIDS2/gcgdata/geneseq/geneseqn/NA2001.DAT:AAS06036	+	107.50	119.78	135.17	3050	!
/SIDS2/gcgdata/geneseq/geneseqn/NA2000.DAT:AAC34308	-	107.00	130.88	32.57	726	!
/SIDS2/gcgdata/geneseq/geneseqn/NA2001.DAT:AAI60744	+	107.00	124.87	70.42	1519	!
/SIDS2/gcgdata/geneseq/geneseqn/NA2001.DAT:AAI60745	+	107.00	124.87	70.42	1519	!
/SIDS2/gcgdata/geneseq/geneseqn/NA1999.DAT:AAV82910	+	106.50	118.10	167.78	3243	!
/SIDS2/gcgdata/geneseq/geneseqn/NA2000.DAT:AAA95033	+	106.50	118.10	167.78	3243	!
/SIDS2/gcgdata/geneseq/geneseqn/NA2000.DAT:AAZ38724	+	106.50	118.10	167.78	3243	!
/SIDS2/gcgdata/geneseq/geneseqn/NA1999.DAT:AAV82909	+	106.50	117.03	192.44	3698	!
/SIDS2/gcgdata/geneseq/geneseqn/NA2000.DAT:AAA95032	+	106.50	117.03	192.44	3698	!
/SIDS2/gcgdata/geneseq/geneseqn/NA2000.DAT:AAZ38723	+	106.50	117.03	192.44	3698	!
/SIDS2/gcgdata/geneseq/geneseqn/NA1999.DAT:AAV82915	+	106.50	116.39	208.99	4002	!
/SIDS2/gcgdata/geneseq/geneseqn/NA2000.DAT:AAA95038	+	106.50	116.39	208.99	4002	!
/SIDS2/gcgdata/geneseq/geneseqn/NA2000.DAT:AAZ38729	+	106.50	116.39	208.99	4002	!
/SIDS2/gcgdata/geneseq/geneseqn/NA1999.DAT:AAV82913	+	106.50	116.36	209.81	4017	!
/SIDS2/gcgdata/geneseq/geneseqn/NA2000.DAT:AAA95036	+	106.50	116.36	209.81	4017	!
/SIDS2/gcgdata/geneseq/geneseqn/NA2000.DAT:AAZ38727	+	106.50	116.36	209.81	4017	!
/SIDS2/gcgdata/geneseq/geneseqn/NA1999.DAT:AAV82912	+	106.50	116.28	211.77	4053	!
/SIDS2/gcgdata/geneseq/geneseqn/NA2000.DAT:AAA95035	+	106.50	116.28	211.77	4053	!
/SIDS2/gcgdata/geneseq/geneseqn/NA2000.DAT:AAZ38726	+	106.50	116.28	211.77	4053	!
/SIDS2/gcgdata/geneseq/geneseqn/NA1999.DAT:AAV82889	+	106.50	116.25	212.59	4068	!
/SIDS2/gcgdata/geneseq/geneseqn/NA2000.DAT:AAA95012	+	106.50	116.25	212.59	4068	!
/SIDS2/gcgdata/geneseq/geneseqn/NA2000.DAT:AAZ38703	+	106.50	116.25	212.59	4068	!
/SIDS2/gcgdata/geneseq/geneseqn/NA1999.DAT:AAV82914	+	106.50	116.24	213.08	4077	!
/SIDS2/gcgdata/geneseq/geneseqn/NA2000.DAT:AAA95037	+	106.50	116.24	213.08	4077	!

seq_name: /SIDS2/gcgdata/geneseq/geneseqn/NA1999.DAT:AAX24995

seq_documentation_block:

```
ID   AAX24995 standard; cDNA; 590 BP.
XX
AC   AAX24995;
XX
DT   05-JUL-1999  (first entry)
XX
DE   Murine Bcl-2 interacting mediator of cell death Bim-EL cDNA.
XX
KW   Bim-EL; Bcl-2 interacting mediator of cell death; apoptosis;
KW   cell cycle; mouse; cancer; autoimmune disease;
KW   degenerative disease; therapy; contraceptive; splice variant;
KW   isoform; ss.
XX
OS   Mus musculus.
XX
PN   WO9914321-A1.
XX
PD   25-MAR-1999.
XX
PF   17-SEP-1998;   98WO-AU00772.
XX
PR   24-SEP-1997;   97AU-0009373.
PR   17-SEP-1997;   97AU-0009263.
XX
PA   (HALL-) HALL INST MEDICAL RES WALTER & ELIZA.
XX
PI   Adams J, Cory S, Huang DCS, O'Connor L, O'Reilly L;
PI   Puthalakath H, Strasser A;
XX
DR   WPI; 1999-244030/20.
DR   P-PSDB; AAW98156.
XX
PT   New isolated member of the Bcl-2 family, Bim used in, e.g. cancer
PT   treatment
XX
PS   Claim 3; Page 96-97; 145pp; English.
XX
CC   The present sequence encodes the extra long form (EL) of murine Bim,
CC   or Bcl-2 interacting mediator of cell death (see AAW98156), a novel
CC   member of the Bcl-2 family that is capable of inducing cell death
CC   (apoptosis) and which acts as a 'death-ligand' for certain members
CC   of the pro-survival Bcl-2 family. Bim is a BH3-only protein, as the
CC   only Bcl-2 homology region which it encompasses is BH3. It is the
CC   only BH3-only protein for which splice variants exist. These
CC   result in the expression of a variety of isoforms, i.e. Bim-S,
CC   Bim-L and Bim-EL (see AAW98154-56). cDNAs encoding these murine Bim
CC   isoforms were obtained from a T lymphoma cDNA library using human
CC   recombinant Bcl-2 protein. The murine Bim gene has been mapped to
CC   chromosome 2 at bands F3-G. Human Bim-L and Bim-EL isoforms have
CC   also been identified (see AAW98157-58). Binding the dynein light
CC   chain was shown to regulate the pro-apoptotic activity of Bim.
CC   Bim-S, the splice variant which does not bind to dynein light
CC   chain, is a much more potent killer than either Bim-L or Bim-EL.
CC   The invention provides variants (see AAW98159-68) of murine and human
CC   Bim-L or Bim-EL that cannot bind, couple or otherwise associate
CC   with a dynein light chain. The identification of Bim permits the
CC   identification and rational design of a range of products for use
CC   in therapy, diagnosis, antibody generation and involving modulation
CC   of physiological cell death. These therapeutic molecules may act
CC   as either antagonists or agonists of Bim's function and will be
CC   useful in cancer, autoimmune or degenerative disease therapy.
CC   Increased Bim expression or Bim activity is useful, e.g. for
CC   treatment or prophylaxis in conditions such as cancer and deletion
CC   of autoreactive lymphocytes in autoimmune disease. Decreased Bim
```

```
CC   expression of Bim activity is useful in regulating inhibition or
CC   prevention of cell death or degeneration such as under cytotoxic
CC   conditions during e.g. gamma-irradiation and chemotherapy or during
CC   HIV/AIDS or other viral infections, ischemia, myocardial infarction,
CC   hypoxia, degenerative diseases or for prolonging the survival of
CC   cells being transplanted for treatment of disease.  Since Bim is
CC   expressed in germ cells, modulating Bim expression or Bim activity
CC   is useful, e.g. as a contraceptive or method of sterilization by
CC   preventing generation of fertile sperm.
XX
SQ   Sequence 590 BP; 137 A; 178 C; 150 G; 125 T; 0 other;


alignment_scores:
            Quality: 1046.00              Length:    196
              Ratio:  5.337                Gaps:      0
 Percent Similarity: 100.000    Percent Identity: 100.000

alignment_block:
 US-09-508-832-6 x AAX24995   ..

 Align seg 1/1  to: AAX24995  from: 1  to: 590

       1 MetAlaLysGlnProSerAspValSerSerGluCysAspArgGluGlyGl 17
         ||||||||||||||||||||||||||||||||||||||||||||||||||
       1 ATGGCCAAGCAACCTTCTGATGTAAGTTCTGAGTGTGACAGAGAAGGTGG 50

      17 yGlnLeuGlnProAlaGluArgProProGlnLeuArgProGlyAlaProT 34
         ||||||||||||||||||||||||||||||||||||||||||||||||||
      51 ACAATTGCAGCCTGCTGAGAGGCCTCCCCAGCTCAGGCCTGGGGCCCCTA 100

      34 hrSerLeuGlnThrGluProGlnGlyAsnProAspGlyGluGlyAspArg 50
       . ||||||||||||||||||||||||||||||||||||||||||||||||||
     101 CCTCCCTACAGACAGAACCGCAAGGTAATCCCGACGGCGAAGGGGACCGC 150

      51 CysProHisGlySerProGlnGlyProLeuAlaProProAlaSerProGl 67
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     151 TGCCCCCACGGCAGCCCTCAGGGCCCGCTGGCCCCACCGGCCAGCCCTGG 200

      67 yProPheAlaThrArgSerProLeuPheIlePheValArgArgSerSerL 84
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     201 CCCTTTTGCTACCAGATCCCCACTTTTCATCTTTGTGAGAAGATCTTCTC 250

      84 euLeuSerArgSerSerSerGlyTyrPheSerPheAspThrAspArgSer 100
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     251 TGCTGTCCCGGTCCTCCAGTGGGTATTTCTCTTTTGACACAGACAGGAGC 300

     101 ProAlaProMetSerCysAspLysSerThrGlnThrProSerProProCy 117
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     301 CCGGCACCCATGAGTTGTGACAAGTCAACACAAACCCCAAGTCCTCCTTG 350

     117 sGlnAlaPheAsnHisTyrLeuSerAlaMetAlaSerIleArgGlnSerG 134
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     351 CCAGGCCTTCAACCACTATCTCAGTGCAATGGCTTCCATACGACAGTCTC 400

     134 lnGluGluProGluAspLeuArgProGluIleArgIleAlaGlnGluLeu 150
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     401 AGGAGGAACCTGAAGATCTGCGCCCGGAGATACGGATTGCACAGGAGCTG 450

     151 ArgArgIleGlyAspGluPheAsnGluThrTyrThrArgArgValPheAl 167
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     451 CGGCGGATCGGAGACGAGTTCAACGAAACTTACACAAGGAGGGTGTTTGC 500

     167 aAsnAspTyrArgGluAlaGluAspHisProGlnMetValIleLeuGlnL 184
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     501 AAATGATTACCGCGAGGCTGAAGACCACCCTCAAATGGTTATCTTACAAC 550

     184 euLeuArgPheIlePheArgLeuValTrpArgArgHis 196
         ||||||||||||||||||||||||||||||||||||
     551 TGTTACGCTTTATCTTCCGTCTGGTATGGAGAAGGCAT 588

seq_name: /SIDS2/gcgdata/geneseq/geneseqn/NA1999.DAT:AAX24997

seq_documentation_block:
ID   AAX24997 standard; cDNA; 596 BP.
XX
AC   AAX24997;
XX
DT   05-JUL-1999  (first entry)
XX
DE   Human Bcl-2 interacting mediator of cell death Bim-EL cDNA.
XX
KW   Bim-EL; Bcl-2 interacting mediator of cell death; apoptosis;
KW   cell cycle; human; cancer; autoimmune disease;
KW   degenerative disease; therapy; contraceptive; splice variant;
KW   isoform; ss.
XX
OS   Homo sapiens.
XX
PN   WO9914321-A1.
XX
PD   25-MAR-1999.
XX
PF   17-SEP-1998;   98WO-AU00772.
XX
PR   24-SEP-1997;   97AU-0009373.
PR   17-SEP-1997;   97AU-0009263.
XX
PA   (HALL-) HALL INST MEDICAL RES WALTER & ELIZA.
XX
PI   Adams J,  Cory S,  Huang DCS,  O'Connor L,  O'Reilly L;
PI   Puthalakath H,  Strasser A;
XX
DR   WPI; 1999-244030/20.
DR   P-PSDB; AAW98158.
XX
PT   New isolated member of the Bcl-2 family, Bim used in, e.g. cancer
PT   treatment
XX
PS   Claim 7; Page 101-102; 145pp; English.
XX
CC   The present sequence encodes the extra long form (EL) of human Bim,
CC   or Bcl-2 interacting mediator of cell death (see AAW98158), a novel
CC   member of the Bcl-2 family that is capable of inducing cell death
CC   (apoptosis) and which acts as a 'death-ligand' for certain members
CC   of the pro-survival Bcl-2 family.  Bim is a BH3-only protein, as the
CC   only Bcl-2 homology region which it encompasses is BH3.  It is the
CC   only BH3-only protein for which splice variants exist.  These
CC   result in the expression of a variety of isoforms, i.e. Bim-S,
CC   Bim-L and Bim-EL.  cDNAs encoding human Bim-L and Bim-EL (see
CC   AAW98158) were isolated from embryo and liver cDNA libraries using
CC   mouse bim cDNA.  Murine Bim-S, Bim-L and Bim-EL isoforms (see
CC   AAW98154-56) are also provided.  The human Bim gene maps to
CC   chromosome 2 at bands 2q12-2q13.  Binding the dynein light
CC   chain was shown to regulate the pro-apoptotic activity of Bim.
CC   Bim-S, the splice variant which does not bind to dynein light
CC   chain, is a much more potent killer than either Bim-L or Bim-EL.
CC   The invention provides variants (see AAW98159-68) of murine and human
CC   Bim-L or Bim-EL that cannot bind, couple or otherwise associate
CC   with a dynein light chain.  The identification of Bim permits the
CC   identification and rational design of a range of products for use
CC   in therapy, diagnosis, antibody generation and involving modulation
CC   of physiological cell death.  These therapeutic molecules may act
CC   as either antagonists or agonists of Bim's function and will be
CC   useful in cancer, autoimmune or degenerative disease therapy.
CC   Increased Bim expression or Bim activity is useful, e.g. for
CC   treatment or prophylaxis in conditions such as cancer and deletion
CC   of autoreactive lymphocytes in autoimmune disease.  Decreased Bim
CC   expression of Bim activity is useful in regulating inhibition or
CC   prevention of cell death or degeneration such as under cytotoxic
CC   conditions during e.g. gamma-irradiation and chemotherapy or during
CC   HIV/AIDS or other viral infections, ischemia, myocardial infarction,
CC   hypoxia, degenerative diseases or for prolonging the survival of
CC   cells being transplanted for treatment of disease.  Since Bim is
CC   expressed in germ cells, modulating Bim expression or Bim activity
CC   is useful, e.g. as a contraceptive or method of sterilization by
```

```
CC   preventing generation of fertile sperm.
XX
SQ   Sequence 596 BP; 145 A; 175 C; 146 G; 130 T; 0 other;


alignment_scores:
            Quality:  908.00              Length:    200
              Ratio:  4.935                Gaps:      2
 Percent Similarity:  92.000   Percent Identity: 86.500

alignment_block:
 US-09-508-832-6 x AAX24997   ..

 Align seg 1/1  to: AAX24997  from: 1  to: 596

       1 MetAlaLysGlnProSerAspValSerSerGluCysAspArgGluGlyGl 17
         ||||||||||||||||||||||||||||||||||||||||||||||||||
       1 ATGGCAAAGCAACCTTCTGATGTAAGTTCTGAGTGTGACCGAGAAGGTAG 50

      17 yGlnLeuGlnProAlaGluArgProProGlnLeuArgProGlyAlaProT 34
         ||||||||||||||||||||||||||||||||||||||||||||||||||
      51 ACAATTGCAGCCTGCGGAGAGGCCTCCCCAGCTCAGACCTGGGGCCCCTA 100

      34 hrSerLeuGlnThrGluProGlnGlyAsnProAsp.............Gly 46
         ||||||||||||||||||||||||||||||||:::             |||
     101 CCTCCCTACAGACAGAGCCACAAGGTAATCCTGAAGGCAATCACGGAGGT 150

      47 GluGlyAspArgCysProHisGlySerProGlnGlyProLeuAlaProPr 63
         |||||||||   ||||||||||||||||||||||||||||||||||||||
     151 GAAGGGGACAGCTGCCCCCACGGCAGCCCTCAGGGCCCGCTGGCCCCACC 200

      63 oAlaSerProGlyProPheAlaThrArgSerProLeuPheIlePheValA 80
         ||||||||||||||||||||||||||||||||||||||||||||||:::|
     201 TGCCAGCCCTGGCCCTTTGCTACCAGATCCCCGCTTTTCATCTTTATGA 250

      80 rgArgSerSerLeuLeuSerArgSerSerSerGlyTyrPheSerPheAsp 96
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     251 GAAGATCCTCCCTGCTGTCTCGATCCTCCAGTGGGTATTTCTCTTTTGAC 300

      97 ThrAspArgSerProAlaProMetSerCysAspLysSerThrGlnThrPr 113
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     301 ACAGACAGGAGCCCAGCACCCATGAGTTGTGACAAATCAACACAAACCCC 350

     113 oSerProProCysGlnAlaPheAsnHisTyrLeuSerAlaMetAlaSerI 130
         ||||||||||||||||||||||||||||||||||||||||||||||||||:
     351 AAGTCCTCCTTGCCAGGCCTTCAACCACTATCTCAGTGCAATGGCTTCCA 400

     130 leArgGlnSerGlnGluGluProGluAspLeuArgProGluIleArgIle 146
         ::||||||:::      ||||||   |||:::|||||||||||   |||
     401 TGAGGCAGGCT......GAACCTGCAGATATGCGCCCAGAGATATGGATC 444

     147 AlaGlnGluLeuArgArgIleGlyAspGluPheAsnGluThrTyrThrAr 163
         ||||||||||||||||||||||||||||||||||      |||:::||
     445 GCCCAAGAGTTGCGGCGTATCGGAGACGAGTTTAACGCTTACTATGCAAG 494

     163 gArgValPheAlaAsnAspTyrArgGluAlaGluAspHisProGlnMetV 180
         ||||||||||   |||:::|||:::   ||||||||||||||:::||||
     495 GAGGGTATTTTTGAATAATTACCAAGCAGCCGAAGACCACCCACGAATGG 544

     180 alIleLeuGlnLeuLeuArgPheIlePheArgLeuValTrpArgArgHis 196
         ||||||||:::|||||||||:::|||   |||||||||||||   |||
     545 TTATCTTACGACTGTTACGTTACATTGTCCGCCTGGTGTGGAGAATGCAT 594

seq_name: /SIDS2/gcgdata/geneseq/geneseqn/NA1999.DAT:AAX24994

seq_documentation_block:
ID   AAX24994 standard; cDNA; 422 BP.
XX
AC   AAX24994;
XX
DT   05-JUL-1999  (first entry)
XX
DE   Murine Bcl-2 interacting mediator of cell death Bim-L cDNA.
XX
KW   Bim-L; Bcl-2 interacting mediator of cell death; apoptosis;
KW   cell cycle; mouse; cancer; autoimmune disease;
KW   degenerative disease; therapy; contraceptive; splice variant;
KW   isoform; ss.
XX
OS   Mus musculus.
XX
PN   WO9914321-A1.
XX
PD   25-MAR-1999.
XX
PF   17-SEP-1998;   98WO-AU00772.
XX
PR   24-SEP-1997;   97AU-0009373.
PR   17-SEP-1997;   97AU-0009263.
XX
PA   (HALL-) HALL INST MEDICAL RES WALTER & ELIZA.
XX
PI   Adams J,  Cory S,  Huang DCS,  O'Connor L,  O'Reilly L;
PI   Puthalakath H,  Strasser A;
XX
DR   WPI; 1999-244030/20.
DR   P-PSDB; AAW98155.
XX
PT   New isolated member of the Bcl-2 family, Bim used in, e.g. cancer
PT   treatment
XX
PS   Claim 3; Page 94-95; 145pp; English.
XX
CC   The present sequence encodes the long form (L) of murine Bim, or
CC   Bcl-2 interacting mediator of cell death (see AAW98155), a novel
CC   member of the Bcl-2 family that is capable of inducing cell death
CC   (apoptosis) and which acts as a 'death-ligand' for certain members
CC   of the pro-survival Bcl-2 family.  Bim is a BH3-only protein, as the
CC   only Bcl-2 homology region which it encompasses is BH3.  It is the
CC   only BH3-only protein for which splice variants exist.  These
CC   result in the expression of a variety of isoforms, i.e. Bim-S,
CC   Bim-L and Bim-EL (see AAW98154-56).  cDNAs encoding these murine Bim
CC   isoforms were obtained from a T lymphoma cDNA library using human
CC   recombinant Bcl-2 protein.  The murine Bim gene has been mapped to
CC   chromosome 2 at bands F3-G.  Human Bim-L and Bim-EL isoforms have
CC   also been identified (see AAW98157-58).  Binding the dynein light
CC   chain was shown to regulate the pro-apoptotic activity of Bim.
CC   Bim-S, the splice variant which does not bind to dynein light
CC   chain, is a much more potent killer than either Bim-L or Bim-EL.
CC   The invention provides variants (see AAW98159-68) of murine and human
CC   Bim-L or Bim-EL that cannot bind, couple or otherwise associate
CC   with a dynein light chain.  The identification of Bim permits the
CC   identification and rational design of a range of products for use
CC   in therapy, diagnosis, antibody generation and involving modulation
CC   of physiological cell death.  These therapeutic molecules may act
CC   as either antagonists or agonists of Bim's function and will be
CC   useful in cancer, autoimmune or degenerative disease therapy.
CC   Increased Bim expression or Bim activity is useful, e.g. for
CC   treatment or prophylaxis in conditions such as cancer and deletion
CC   of autoreactive lymphocytes in autoimmune disease.  Decreased Bim
CC   expression of Bim activity is useful in regulating inhibition or
CC   prevention of cell death or degeneration such as under cytotoxic
CC   conditions during e.g. gamma-irradiation and chemotherapy or during
CC   HIV/AIDS or other viral infections, ischemia, myocardial infarction,
CC   hypoxia, degenerative diseases or for prolonging the survival of
CC   cells being transplanted for treatment of disease.  Since Bim is
CC   expressed in germ cells, modulating Bim expression or Bim activity
CC   is useful, e.g. as a contraceptive or method of sterilization by
CC   preventing generation of fertile sperm.
XX
SQ   Sequence 422 BP; 112 A; 116 C; 109 G; 85 T; 0 other;


alignment_scores:
            Quality:  704.00              Length:    196
              Ratio:  5.029                Gaps:      1
```

```
 Percent Similarity:  71.429   Percent Identity: 71.429

alignment_block:
 US-09-508-832-6 x AAX24994    ..

 Align seg 1/1  to: AAX24994  from: 1  to: 422

       1 MetAlaLysGlnProSerAspValSerSerGluCysAspArgGluGlyGl 17
         ||||||||||||||||||||||||||||||||||||||||||||||||||
       1 ATGGCCAAGCAACCTTCTGATGTAAGTTCTGAGTGTGACAGAGAAGGTGG 50

      17 yGlnLeuGlnProAlaGluArgProProGlnLeuArgProGlyAlaProT 34
         ||||||||||||||||||||||||||||||||||||||||||||||||||
      51 ACAATTGCAGCCTGCTGAGAGGCCTCCCCAGCTCAGGCCTGGGGCCCCTA 100

      34 hrSerLeuGlnThrGluProGlnGlyAsnProAspGlyGluGlyAspArg 50
         ||||||||||||||||||||||
     101 CCTCCCTACAGACAGAACCGCAA..........................  123

      51 CysProHisGlySerProGlnGlyProLeuAlaProProAlaSerProGl 67

     123 ..................................................  123

      67 yProPheAlaThrArgSerProLeuPheIlePheValArgArgSerSerL 84

     123 ..................................................  123

      84 euLeuSerArgSerSerSerGlyTyrPheSerPheAspThrAspArgSer 100
                                                 |||||||||
     124 ....................................GACAGGAGC 132

     101 ProAlaProMetSerCysAspLysSerThrGlnThrProSerProProCy 117
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     133 CCGGCACCCATGAGTTGTGACAAGTCAACACAAACCCCAAGTCCTCCTTG 182

     117 sGlnAlaPheAsnHisTyrLeuSerAlaMetAlaSerIleArgGlnSerG 134
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     183 CCAGGCCTTCAACCACTATCTCAGTGCAATGGCTTCCATACGACAGTCTC 232

     134 lnGluGluProGluAspLeuArgProGluIleArgIleAlaGlnGluLeu 150
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     233 AGGAGGAACCTGAAGATCTGCGCCCGGAGATACGGATTGCACAGGAGCTG 282

     151 ArgArgIleGlyAspGluPheAsnGluThrTyrThrArgArgValPheAl 167
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     283 CGGCGGATCGGAGACGAGTTCAACGAAACTTACACAAGGAGGGTGTTTGC 332

     167 aAsnAspTyrArgGluAlaGluAspHisProGlnMetValIleLeuGlnL 184
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     333 AAATGATTACCGCGAGGCTGAAGACCACCCTCAAATGGTTATCTTACAAC 382

     184 euLeuArgPheIlePheArgLeuValTrpArgArgHis 196
         ||||||||||||||||||||||||||||||||||||
     383 TGTTACGCTTTATCTTCCGTCTGGTATGGAGAAGGCAT 420

seq_name: /SIDS2/gcgdata/geneseq/geneseqn/NA1999.DAT:AAX24996

seq_documentation_block:
ID   AAX24996 standard; cDNA; 416 BP.
XX
AC   AAX24996;
XX
DT   05-JUL-1999  (first entry)
XX
DE   Human Bcl-2 interacting mediator of cell death Bim-L cDNA.
XX
KW   Bim-L; Bcl-2 interacting mediator of cell death; apoptosis;
KW   cell cycle; human; cancer; autoimmune disease;
KW   degenerative disease; therapy; contraceptive; splice variant;
KW   isoform; ss.
XX
OS   Homo sapiens.
XX
PN   WO9914321-A1.
XX
PD   25-MAR-1999.
XX
PF   17-SEP-1998;   98WO-AU00772.
XX
PR   24-SEP-1997;   97AU-0009373.
PR   17-SEP-1997;   97AU-0009263.
XX
PA   (HALL-) HALL INST MEDICAL RES WALTER & ELIZA.
XX
PI   Adams J,  Cory S,  Huang DCS,  O'Connor L,  O'Reilly L;
PI   Puthalakath H,  Strasser A;
XX
DR   WPI; 1999-244030/20.
DR   P-PSDB; AAW98157.
XX
PT   New isolated member of the Bcl-2 family, Bim used in, e.g. cancer
PT   treatment
XX
PS   Claim 7; Page 99-100; 145pp; English.
XX
CC   The present sequence encodes the long form (L) of human Bim, or
CC   Bcl-2 interacting mediator of cell death (see AAW98157), a novel
CC   member of the Bcl-2 family that is capable of inducing cell death
CC   (apoptosis) and which acts as a 'death-ligand' for certain members
CC   of the pro-survival Bcl-2 family.  Bim is a BH3-only protein, as the
CC   only Bcl-2 homology region which it encompasses is BH3.  It is the
CC   only BH3-only protein for which splice variants exist.  These
CC   result in the expression of a variety of isoforms, i.e. Bim-S,
CC   Bim-L and Bim-EL.  cDNAs encoding human Bim-L and Bim-EL (see
CC   AAW98158) were isolated from embryo and liver cDNA libraries using
CC   mouse bim cDNA.  Murine Bim-S, Bim-L and Bim-EL isoforms (see
CC   AAW98154-56) are also provided.  The human Bim gene maps to
CC   chromosome 2 at bands 2q12-2q13.  Binding the dynein light
CC   chain was shown to regulate the pro-apoptotic activity of Bim.
CC   Bim-S, the splice variant which does not bind to dynein light
CC   chain, is a much more potent killer than either Bim-L or Bim-EL.
CC   The invention provides variants (see AAW98159-68) of murine and human
CC   Bim-L or Bim-EL that cannot bind, couple or otherwise associate
CC   with a dynein light chain.  The identification of Bim permits the
CC   identification and rational design of a range of products for use
CC   in therapy, diagnosis, antibody generation and involving modulation
CC   of physiological cell death.  These therapeutic molecules may act
CC   as either antagonists or agonists of Bim's function and will be
CC   useful in cancer, autoimmune or degenerative disease therapy.
CC   Increased Bim expression or Bim activity is useful, e.g. for
CC   treatment or prophylaxis in conditions such as cancer and deletion
CC   of autoreactive lymphocytes in autoimmune disease.  Decreased Bim
CC   expression of Bim activity is useful in regulating inhibition or
CC   prevention of cell death or degeneration such as under cytotoxic
CC   conditions during e.g. gamma-irradiation and chemotherapy or during
CC   HIV/AIDS or other viral infections, ischemia, myocardial infarction,
CC   hypoxia, degenerative diseases or for prolonging the survival of
CC   cells being transplanted for treatment of disease.  Since Bim is
CC   expressed in germ cells, modulating Bim expression or Bim activity
CC   is useful, e.g. as a contraceptive or method of sterilization by
CC   preventing generation of fertile sperm.
XX
SQ   Sequence 416 BP; 113 A; 113 C; 103 G; 87 T; 0 other;


alignment_scores:
           Quality:  591.00              Length:    196
             Ratio:  4.581                 Gaps:      2
 Percent Similarity:  65.816   Percent Identity: 61.224

alignment_block:
 US-09-508-832-6 x AAX24996    ..

 Align seg 1/1  to: AAX24996  from: 1  to: 416

       1 MetAlaLysGlnProSerAspValSerSerGluCysAspArgGluGlyGl 17
```

```
       ||||||||||||||||||||||||||||||||||||||||||||||||||
     1 ATGGCAAAGCAACCTTCTGATGTAAGTTCTGAGTGTGACCGAGAAGGTAG 50

    17 yGlnLeuGlnProAlaGluArgProProGlnLeuArgProGlyAlaProT 34
       ||||||||||||||||||||||||||||||||||||||||||||||||||
    51 ACAATTGCAGCCTGCGGAGAGGCCTCCCCAGCTCAGACCTGGGGCCCCTA 100

    34 hrSerLeuGlnThrGluProGlnGlyAsnProAspGlyGluGlyAspArg 50
       |||||||||||||||||||||
   101 CCTCCCTACAGACAGAGCCACAA........................... 123

    51 CysProHisGlySerProGlnGlyProLeuAlaProProAlaSerProGl 67

   123 .................................................. 123

    67 yProPheAlaThrArgSerProLeuPheIlePheValArgArgSerSerL 84

   123 .................................................. 123

    84 euLeuSerArgSerSerSerGlyTyrPheSerPheAspThrAspArgSer 100
                                                 |||||||||
   124 .........................................GACAGGAGC 132

   101 ProAlaProMetSerCysAspLysSerThrGlnThrProSerProProCy 117
       ||||||||||||||||||||||||||||||||||||||||||||||||||
   133 CCAGCACCCATGAGTTGTGACAAATCAACACAAACCCCAAGTCCTCCTTG 182

   117 sGlnAlaPheAsnHisTyrLeuSerAlaMetAlaSerIleArgGlnSerG 134
       ||||||||||||||||||||||||||||||||||||:::||||||:::
   183 CCAGGCCTTCAACCACTATCTCAGTGCAATGGCTTCCATGAGGCAGGCT. 231

   134 lnGluGluProGluAspLeuArgProGluIleArgIleAlaGlnGluLeu 150
            ||||||   |||:::||||||||||||   ||||||||||||||
   232 .....GAACCTGCAGATATGCGCCCAGAGATATGGATCGCCCAAGAGTTG 276

   151 ArgArgIleGlyAspGluPheAsnGluThrTyrThrArgArgValPheAl 167
       ||||||||||||||||||||||          |||:::||||||||||
   277 CGGCGTATCGGAGACGAGTTTAACGCTTACTATGCAAGGAGGGTATTTTT 326

   167 aAsnAspTyrArgGluAlaGluAspHisProGlnMetValIleLeuGlnL 184
        |||:::|||:::   ||||||||||||||||:::||||||||||||:::|
   327 GAATAATTACCAAGCAGCCGAAGACCACCCACGAATGGTTATCTTACGAC 376

   184 euLeuArgPheIlePheArgLeuValTrpArgArgHis 196
       |||||||:::|||   ||||||||||||||    |||
   377 TGTTACGTTACATTGTCCGCCTGGTGTGGAGAATGCAT 414
```

seq_name: /SIDS2/gcgdata/geneseq/geneseqn/NA1999.DAT:AAX24993

```
seq_documentation_block:
ID   AAX24993 standard; cDNA; 332 BP.
XX
AC   AAX24993;
XX
DT   05-JUL-1999  (first entry)
XX
DE   Murine Bcl-2 interacting mediator of cell death Bim-S cDNA.
XX
KW   Bim-S; Bcl-2 interacting mediator of cell death; apoptosis;
KW   cell cycle; mouse; cancer; autoimmune disease;
KW   degenerative disease; therapy; contraceptive; splice variant;
KW   isoform; ss.
XX
OS   Mus musculus.
XX
PN   WO9914321-A1.
XX
PD   25-MAR-1999.
XX
PF   17-SEP-1998;   98WO-AU00772.
XX
PR   24-SEP-1997;   97AU-0009373.
PR   17-SEP-1997;   97AU-0009263.
XX
PA   (HALL-) HALL INST MEDICAL RES WALTER & ELIZA.
XX
PI   Adams J,  Cory S,  Huang DCS,  O'Connor L,  O'Reilly L;
PI   Puthalakath H,  Strasser A;
XX
DR   WPI; 1999-244030/20.
DR   P-PSDB; AAW98154.
XX
PT   New isolated member of the Bcl-2 family, Bim used in, e.g. cancer
PT   treatment
XX
PS   Claim 3; Page 92; 145pp; English.
XX
CC   The present sequence encodes the short form (S) of murine Bim, or
CC   Bcl-2 interacting mediator of cell death (see AAW98154), a novel
CC   member of the Bcl-2 family that is capable of inducing cell death
CC   (apoptosis) and which acts as a 'death-ligand' for certain members
CC   of the pro-survival Bcl-2 family.  Bim is a BH3-only protein, as the
CC   only Bcl-2 homology region which it encompasses is BH3.  It is the
CC   only BH3-only protein for which splice variants exist.  These
CC   result in the expression of a variety of isoforms, i.e. Bim-S,
CC   Bim-L and Bim-EL (see AAW98154-56).  cDNAs encoding these murine Bim
CC   isoforms were obtained from a T lymphoma cDNA library using human
CC   recombinant Bcl-2 protein.  The murine Bim gene has been mapped to
CC   chromosome 2 at bands F3-G.  Human Bim-L and Bim-EL isoforms have
CC   also been identified (see AAW98157-58).  Binding the dynein light
CC   chain was shown to regulate the pro-apoptotic activity of Bim.
CC   Bim-S, the splice variant which does not bind to dynein light
CC   chain, is a much more potent killer than either Bim-L or Bim-EL.
CC   The invention provides variants (see AAW98159-68) of murine and human
CC   Bim-L or Bim-EL that cannot bind, couple or otherwise associate
CC   with a dynein light chain.  The identification of Bim permits the
CC   identification and rational design of a range of products for use
CC   in therapy, diagnosis, antibody generation and involving modulation
CC   of physiological cell death.  These therapeutic molecules may act
CC   as either antagonists or agonists of Bim's function and will be
CC   useful in cancer, autoimmune or degenerative disease therapy.
CC   Increased Bim expression or Bim activity is useful, e.g. for
CC   treatment or prophylaxis in conditions such as cancer and deletion
CC   of autoreactive lymphocytes in autoimmune disease.  Decreased Bim
CC   expression of Bim activity is useful in regulating inhibition or
CC   prevention of cell death or degeneration such as under cytotoxic
CC   conditions during e.g. gamma-irradiation and chemotherapy or during
CC   HIV/AIDS or other viral infections, ischemia, myocardial infarction,
CC   hypoxia, degenerative diseases or for prolonging the survival of
CC   cells being transplanted for treatment of disease.  Since Bim is
CC   expressed in germ cells, modulating Bim expression or Bim activity
CC   is useful, e.g. as a contraceptive or method of sterilization by
CC   preventing generation of fertile sperm.
XX
SQ   Sequence 332 BP; 87 A; 85 C; 91 G; 69 T; 0 other;
```

```
alignment_scores:
          Quality:  521.00            Length:    196
            Ratio:  4.736               Gaps:      1
 Percent Similarity:  56.122   Percent Identity: 56.122

alignment_block:
 US-09-508-832-6 x AAX24993

 Align seg 1/1  to: AAX24993  from: 1  to: 332

     1 MetAlaLysGlnProSerAspValSerSerGluCysAspArgGluGlyGl 17
       ||||||||||||||||||||||||||||||||||||||||||||||||||
     1 ATGGCCAAGCAACCTTCTGATGTAAGTTCTGAGTGTGACAGAGAAGGTGG 50

    17 yGlnLeuGlnProAlaGluArgProProGlnLeuArgProGlyAlaProT 34
       ||||||||||||||||||||||||||||||||||||||||||||||||||
    51 ACAATTGCAGCCTGCTGAGAGGCCTCCCCAGCTCAGGCCTGGGGCCCCTA 100

    34 hrSerLeuGlnThrGluProGlnGlyAsnProAspGlyGluGlyAspArg 50
```

```
     ||||||||||||||||||||||
 101 CCTCCCTACAGACAGAACCGCAA..................................... 123

  51 CysProHisGlySerProGlnGlyProLeuAlaProProAlaSerProGl 67

 123 ..... 123

  67 yProPheAlaThrArgSerProLeuPheIlePheValArgArgSerSerL 84

 123 ..... 123

  84 euLeuSerArgSerSerSerGlyTyrPheSerPheAspThrAspArgSer 100

 123 ..... 123

 101 ProAlaProMetSerCysAspLysSerThrGlnThrProSerProProCy 117

 123 ..... 123

 117 sGlnAlaPheAsnHisTyrLeuSerAlaMetAlaSerIleArgGlnSerG 134
                                  ||||||||||||||||||||
 124 ..........................GCTTCCATACGACACGTCTC 142

 134 lnGluGluProGluAspLeuArgProGluIleArgIleAlaGlnGluLeu 150
     ||||||||||||||||||||||||||||||||||||||||||||||||||||
 143 AGGAGGAACCTGAAGATCTGCGCCCGGAGATACGGATTGCACAGGAGCTG 192

 151 ArgArgIleGlyAspGluPheAsnGluThrTyrThrArgArgValPheAl 167
     ||||||||||||||||||||||||||||||||||||||||||||||||||||
 193 CGGCGGATCGGAGACGAGTTCAACGAAACTTACACAAGGAGGGTGTTTGC 242

 167 aAsnAspTyrArgGluAlaGluAspHisProGlnMetValIleLeuGlnL 184
     ||||||||||||||||||||||||||||||||||||||||||||||||||||
 243 AAATGATTACCGCGAGGCTGAAGACCACCCTCAAATGGTTATCTTACAAC 292

 184 euLeuArgPheIlePheArgLeuValTrpArgArgHis 196
     ||||||||||||||||||||||||||||||||||||
 293 TGTTACGCTTTATCTTCCGTCTGGTATGGAGAAGGCAT 330
```

seq_name: /SIDS2/gcgdata/geneseq/geneseqn/NA1998.DAT:AAV35620

seq_documentation_block:

ID AAV35620 standard; DNA; 32367 BP.
XX
AC AAV35620;
XX
DT 07-SEP-1998 (first entry)
XX
DE Human SHOX (short stature homeobox containing gene) gene sequence.
XX
KW Homeobox domain; human growth gene; growth regulation; growth defect;
KW turner's syndrome; short stature homeobox containing gene; SHOXa;
KW SHOX; bone disease; osteoporosis; calcium regulation; short stature;
KW transcription factor A; ss.
XX
OS Homo sapiens.
XX
PN WO9814568-A1.
XX
PD 09-APR-1998.
XX
PF 29-SEP-1997; 97WO-EP05355.
XX
PR 16-JAN-1997; 97EP-0100583.
PR 01-OCT-1996; 96US-0027633.
XX
PA (RAPP/) RAPPOLD-HOERBRAND G.
XX
PI Rao E, Rappold-hoerbrand G;
XX
DR WPI; 1998-271719/24.
XX
PT New human growth genes - used to develop products for the diagnosis
PT and treatment of human growth defects such as short stature, e.g.
PT Turner's syndrome
XX
PS Claim 19; Pages 51-67; 84pp; English.
XX
CC This is the human SHOX gene sequence containing the PAR1 region. The
CC gene region corresponding to short stature has been identified as a
CC region of approximately 500 kb in the PAR1 region of the X and Y
CC chromosomes. Three genes in this region have been identified as
CC candidates for the short stature gene. These genes were designated SHOX
CC (also referred to as SHOX93 or HOX93), pET92 and SHOT (SHOX-like homeobox
CC gene on chromosome three). The SHOX gene has two separate splicing sites
CC resulting in two variations SHOXa and SHOXb. The specification provides
CC sequences of SHOX (short stature homeobox-containing) genes SHOX ET92,
CC SHOXa, SHOXb, SHOT and exons of the SHOX genes as shown in AAV35610 to
CC AAV35621 and protein sequences of the human growth protein transcription
CC factor SHOXa, SHOXb and SHOT as shown AAW60573 to AAW60575. The novel
CC genes are responsible for human growth. Defects in the genes can cause
CC short stature, e.g. Turner's syndrome. The products can be used to
CC develop agents for the treatment of short stature or other human growth
CC disorders. The products can also be used for providing a mitogenic effect
CC on cells, e.g. for the treatment of bone diseases such as osteoporosis
CC and diseases involved with disturbance in the bone calcium regulation.
XX
SQ Sequence 32367 BP; 7627 A; 8130 C; 8564 G; 8043 T; 3 other;

alignment_scores:

```
           Quality:  120.50            Length:    159
             Ratio:  1.470               Gaps:     11
 Percent Similarity:  51.572   Percent Identity: 33.962
```

alignment_block:

US-09-508-832-6 x AAV35620/rev ..

Align_seg 1/1 to reverse of: AAV35620 from: 1 to: 32367

```
      5 ProSerAspValSerSerGluCysAspArgGluGlyGlyGlnLeuGlnPr 21
        |||:::  :::           |||  :::|||:::         ||
  14059 CCCCAGCCTCTCCCCAGGACACCAGATCCCCAAGGCTCTCCCAGGACACC 14010

     21 oAlaGluArgProProGlnLeuArgProGlyAlaProThrSerLeuGlnT 38
        |:::       ||||||     ||||||:::|||
  14009 AGGT.........CCCCAAGCCTCTCCCGGGACACCA............. 13982

     38 hrGluProGlnGlyAsnProAspGlyGluGlyAspArgCysProHisGly 54
        ||||||:::::|||  :::  |||       |||:::|||
  13981 ..GGTCCCCAAGCCTCTCCCGGGTCACTAGGT.........CCCCAAGGC 13943

     55 SerProGln......GlyProLeuAlaProProAlaSerProGlyProPh 69
        :::|||:::      ||||||  |||  |||  ||||||||||||
  13942 ACTCCCGAGACACCAGGTCCCCAAGCCTCTCCAAGGTCACCAGGTCCCCA 13893

     69 eAla..........ThrArgSerProLeuPheIlePheValArgArgS 82
        :::           ||||||||||||  :::      |||  ||||
  13892 AGGCTCTCCCCAGGACACCAGGTCCCCAAGGCTCTCCCGGGTCACCAGGT 13843

     82 erSerLeuLeuSerArg...SerSerSerGlyTyrPheSerPheAspThr 97
        ||    ||||||:::  :::  |||      :::|||:::||||||
  13842 CCCCAGGCCTCTCCAAGGTCACCAGGTCCCCAAGCCTCTCCTGGGACACC 13793

     98 ..............AspArg.SerProAlaProMetSerCysAspLysS 109
                    :::||| :::|||:::|||::::::         |
  13792 AGGTGTTCAAGCCTCTCCCGGGACACCAAGTCCCCAAGCC.........T 13752

    109 erThrGlnThrPro.....................SerProProCysGln 118
        ||     ||||||                    ||||||      |||
  13751 CTCCCGGGACACCAGGTCCCCAAGCCTCTCCCGGCTCACCAGATCTCCAA 13702

    119 AlaPheAsnHisTyrLeuSerAlaMetAlaSerIleArgGln......Se 133
        |||    :::      :::  :::|||||   |||
  13701 GCCTCTCCCAGGACACCAGGTCCCCAAGCCTCTCCCAGGACACCAGGTCC 13652
```

```
    133 rGlnGluGluProGluAspLeuArg 141
        |||   :::|||:::|||   |||
  13651 CCAAGCCTCTCCCCAGGACACCAGG 13627

seq_name: /SIDS2/gcgdata/geneseq/geneseqn/NA2001.DAT:AAH17812

seq_documentation_block:
ID   AAH17812 standard; cDNA; 2405 BP.
XX
AC   AAH17812;
XX
DT   26-JUN-2001  (first entry)
XX
DE   Human cDNA sequence SEQ ID NO:17475.
XX
KW   Human; primer; detection; diagnosis; antisense therapy; gene therapy; ss.
XX
OS   Homo sapiens.
XX
PN   EP1074617-A2.
XX
PD   07-FEB-2001.
XX
PF   28-JUL-2000; 2000EP-0116126.
XX
PR   29-JUL-1999;   99JP-0248036.
PR   27-AUG-1999;   99JP-0300253.
PR   11-JAN-2000; 2000JP-0118776.
PR   02-MAY-2000; 2000JP-0183767.
PR   09-JUN-2000; 2000JP-0241899.
XX
PA   (HELI-) HELIX RES INST.
XX
PI   Ota T,  Isogai T,  Nishikawa T,  Hayashi K,  Saito K,  Yamamoto J;
PI   Ishii S,  Sugiyama T,  Wakamatsu A,  Nagai K,  Otsuki T;
XX
DR   WPI; 2001-318749/34.
XX
PT   Primer sets for synthesizing polynucleotides, particularly the 5602
PT   full-length cDNAs defined in the specification, and for the detection
PT   and/or diagnosis of the abnormality of the proteins encoded by the
PT   full-length cDNAs -
XX
PS   Claim 8; SEQ ID 17475; 2537pp + CD ROM; English.
XX
CC   The present invention describes primer sets for synthesising 5602
CC   full-length cDNAs defined in the specification. Where a primer set
CC   comprises: (a) an oligo-dT primer and an oligonucleotide complementary
CC   to the complementary strand of a polynucleotide which comprises one of
CC   the 5602 nucleotide sequences defined in the specification, where the
CC   oligonucleotide comprises at least 15 nucleotides; or (b) a combination
CC   of an oligonucleotide comprising a sequence complementary to the
CC   complementary strand of a polynucleotide which comprises a 5'-end
CC   sequence and an oligonucleotide comprising a sequence complementary to a
CC   polynucleotide which comprises a 3'-end sequence, where the
CC   oligonucleotide comprises at least 15 nucleotides and the combination of
CC   the 5'-end sequence/3'-end sequence is selected from those defined in
CC   the specification. The primer sets can be used in antisense therapy and
CC   in gene therapy. The primers are useful for synthesising polynucleotides,
CC   particularly full-length cDNAs. The primers are also useful for the
CC   detection and/or diagnosis of the abnormality of the proteins encoded by
CC   the full-length cDNAs. The primers allow obtaining of the full-length
CC   cDNAs easily without any specialised methods. AAH03166 to AAH13628 and
CC   AAH13633 to AAH18742 represent human cDNA sequences; AAB92446 to
CC   AAB95893 represent human amino acid sequences; and AAH13629 to AAH13632
CC   represent oligonucleotides, all of which are used in the exemplification
CC   of the present invention.
XX
SQ   Sequence 2405 BP; 508 A; 784 C; 739 G; 374 T; 0 other;


alignment_scores:
            Quality:  118.50              Length:    198
            Ratio:  1.445                 Gaps:      5
Percent Similarity:  41.414   Percent Identity: 24.242

alignment_block:
US-09-508-832-6 x AAH17812   ..

 Align seg 1/1  to: AAH17812  from: 1  to: 2405

      2 AlaLysGlnProSerAspValSerSerGluCysAspArgGluGlyGlyGl 18
        |||  ::::  |||||||:::   :::::::|||         |||:::
    498 GCGCCCGAGGTCTCCGACGCGCGGAAACGCTGCTTCGCCCTAGGCGCAGT 547

     18 nLeuGlnProAlaGluArgProProGlnLeuArgProGlyAlaProThrS 35
                      ||||||         |||   ||||||
    548 GGGGCCAGGACTCCCCACGCCGCCGCCGCCGCCGCCTCCTGCGCCC.... 593

     35 erLeuGlnThrGluProGlnGlyAsnProAspGlyGluGlyAspArgCys 51
           |||::::::      |||:::|||:::::::::      |||
    594 .....CAGAGCCAGGCACCTGGGGGCCCAGAGGCACAGCCTTTCCGGGAG 638

     52 ProHisGlySerProGlnGlyProLeu...AlaProProAlaSerProGl 67
        |||         |||:::      |||   ||||||||||||   |||::
    639 CCGGGTCCGCGTCCTCGCATCTTGCTGTGCGCACCGCCCGCGCGCCCCGC 688

     67 yProPheAlaThrArgSerProLeuPheIlePheValArgArgSerSerL 84
        :|||  ||||
    689 GCCGTCAGCA..................................... 698

     84 euLeuSerArgSerSerSerGlyTyrPheSerPheAspThrAspArgSer 100

    699 .............................................CCC 701

    101 ProAlaProMetSerCysAspLysSerThrGlnThrProSerProProCy 117
        |||||||||   :::      :::||||||      |||:::||||||
    702 CCAGCACCGCCAGCGCCCCCGGAGTCCACTGTGCGCCCTGCGCCCCCGAC 751

    117 sGln.................................AlaPheAsnH 122
        :::                                 :::|||:
    752 GCGCCCCGGGGAAAGTTCCTACTCGTCAATTTCACACGTAATTTACAATA 801

    122 isTyrLeuSerAlaMetAlaSerIleArgGlnSerGlnGluGluProGlu 138
        :::::   ::::::   ||||||   |||:::         |||
    802 ACCACCAGGATTCCTCCGCGTCGCCTAGGAAACGACCGGGCGAAGCGACT 851

    139 AspLeuArgProGluIleArgIleAlaGlnGluLeuArgArgIleGlyAs 155
                  ||||||:::      |||:::
    852 GCCGCCTCCTCCGAGATCAAAGCCCTGCAGCAG.............. 884

    155 pGluPheAsnGluThrTyrThrArgArgValPheAlaAsnAspTyrArgG 172
                        |||||||||::::::||||||      ::::
    885 ..................ACCCGGAGGCTCCTGGCGAACGCCAGGGAGC 915

    172 luAlaGluAspHisProGlnMetValIleLeuGlnLeuLeuArg 186
        ::::::::   |||         :::   ::::::   ||||||
    916 GGACGCGGGTGCACACCATCAGCGCAGCCTTCGAGGCGCTCAGG 959

seq_name: /SIDS2/gcgdata/geneseq/geneseqn/NA2000.DAT:AAZ61771

seq_documentation_block:
ID   AAZ61771 standard; cDNA; 1421 BP.
XX
AC   AAZ61771;
XX
DT   27-MAR-2000  (first entry)
XX
DE   cDNA encoding mouse skin cell transmembrane protein, SEQ ID NO:244.
XX
KW   Skin; dermal papilla; keratinocyte; neonatal foreskin fibroblast;
KW   embryonic skin cell; keratinocyte stem cell; transit amplifying cell;
KW   secreted; transmembrane; inflammation; cancer; neurological disease;
KW   angiogenesis; tumour vascularisation; growth disorder;
KW   developmental disorder; skin wound; hair follicle disorder;
KW   anti-inflammatory; cytostatic; neuroprotective; vulnery; ss.
```

```
XX
OS   Mus sp.
XX
PN   WO9955865-A1.
XX
PD   04-NOV-1999.
XX
PF   29-APR-1999;   99WO-NZ00051.
XX
PR   29-APR-1998;   98US-0069726.
PR   09-NOV-1998;   98US-0188930.
XX
PA   (GENE-) GENESIS RES & DEV CORP LTD.
XX
PI   Strachan L,  Sleeman M,  Watson JD,  Onrust R,  Kumble A,  Murison JG;
XX
DR   WPI; 2000-072177/06.
XX
PT   Novel polynucleotides useful for the treatment of various conditions
PT   including wounds and cancer -
XX
PS   Claim 1; Page 153; 235pp; English.
XX
CC   The invention relates to novel nucleic acid sequences derived from rat
CC   dermal papilla, human keratinocytes and neonatal foreskin fibroblasts,
CC   and mouse embryonic skin, keratinocyte stem cells and transit amplifying
CC   cells. Polypeptides of the invention may be used to treat inflammation,
CC   cancer and neurological diseases. The proteins may be used to stimulate
CC   the growth and motility of keratinocytes, to inhibit the growth of
CC   cancer cells, to modulate angiogenesis and tumour vascularisation, to
CC   modulate skin inflammation, to modulate epithelial cell growth and to
CC   inhibit binding of HIV-1 to leukocytes. The invention may also be used
CC   to treat growth and developmental defects, skin wounds and hair follicle
CC   disorders. Sequences AAZ61606-Z61832 represent cDNA sequences derived
CC   from several mouse, rat or human skin cell types. Sequences
CC   AAZ61606-Z61649, AAZ61725-Z61765, AAZ61802-Z61811 and AAZ61826 encode
CC   proteins with an N-terminal signal sequence, indicating that the proteins
CC   are secreted. Sequences AAZ61650-Z61668, AAZ61766-Z61780, AAZ61812-Z61817
CC   and AAZ61827-Z61829 encode proteins with one or more putative
CC   transmembrane domains.
XX
SQ   Sequence 1421 BP; 290 A; 435 C; 429 G; 265 T; 2 other;


alignment_scores:
            Quality:  112.50              Length:    209
              Ratio:  1.103                Gaps:     13
 Percent Similarity:  48.804   Percent Identity: 28.230

alignment_block:
 US-09-508-832-6 x AAZ61771   ..

 Align seg 1/1  to: AAZ61771  from: 1  to: 1421

       3 LysGlnProSerAspValSerSerGluCysAspArgGlu......GlyGl 17
         |||      |||:::   ::::::|||   ::::::|||      :::||
     738 AAGGGTGGGTCGGAGGAGGAGGAGGAGAAGGAAGAGGAGGTGTTCAGTGG 787

      17 yGlnLeuGlnProAlaGluArgPro............ProGlnLeuArgP 30
         ||||   :::|||   |||   |||                :::   :::|
     788 CCAGGAGGAGCCCCGGGAAGCCCCAGTATGTGGGGTCACTGAAGAGAAGC 837

      30 roGlyAlaPro...ThrSerLeuGlnThrGluProGlnGlyAsnPro... 44
         ||   :::|||      :::   :::::::|||   :::|||   |||
     838 CGGAGGTCCCTGACGAGACAGCCTCAGCAGAGGCTGAAGGGGTTCCCGCA 887

      45 ......AspGlyGluGlyAspArgCysProHisGlySer.......... 55
               :::|||:::|||:::      |||:::||||||
     888 GCCAGCGAGGGCCAAGGGGAA......CCAGAAGGGTCTTTCTCCTTAGC 931

      56 .......ProGlnGlyProLeuAlaPro.ProAlaSerProGlyProPhe 69
               |||||||||      :::|||  |||   :::|||   |||
     932 CCAGGAACCCCAGGGAGCAGCTGGTCCTTCCGAAAGGTCCTGTGCCT... 978

      70 AlaThrArgSerProLeu.............................. 75
         ||||||:::|||||||||
     979 GCAACAGAATCTCCCCTAATGTGTAACAGGCCCCAGAACTGTGAGGCCTG 1028

      76 .PheIlePheValArgArgSerSerLeuLeuSerArgSerSerSerGlyT 92
          ::::::      |||||||||         ||||||::::::
    1029 ACTCTTGGGTCCTCGAAGGTCACCTCCTTGGTCAAGAAAGGCA....... 1071

      92 yrPheSerPheAspThrAspArgSerPro..................... 101
           ||||||||||||         :::|||
    1072 ..TTCAGCTTTGACTGCTTCTTGACACCCTGCCTTGGCCATTGTGGGTGC 1119

     102 .....:.....................AlaProMetSerCysAspLysSe 109
                                     |||   :::   |||      ::
    1120 CAATCCTGACCCTGAATGGGCAAAGCTGCTGGCCTCTGGTGTACCCCAGG 1169

     109 rThrGlnThrProSerProProCysGlnAlaPheAsnHisTyrLeuSerA 126
         ::::   |||||||||       :::          :::|||   |||:::|
    1170 AAACACCACCCCAAGTTCCAGCGCCCTTAATGACTCTCACATCCTGGGGG 1219

     126 laMetAlaSerIleArg..............GlnSerGlnGluGluPro 137
         ||                            ::::::|||   :::|||
    1220 CTTCACCCCGAAGCACCACTTTTCTGGAAGGGGAAGGTCAGACACATCCC 1269

     138 GluAspLeuArgProGluIleArgIleAlaGlnGluLeuArgArgIleGl 154
            :::|||:::|||:::            |||:::   |||      ||
    1270 ...AGTTTGGAGCCGCAATGAGGCAGTCCTCAGAACAGAAGGGGAACAGG 1316

     154 yAspGluPheAsnGluThrTyrThr 162
         |::::::::::   |||||||||
    1317 CCAGAGGCTGACTGTGACATACACA 1341

seq_name: /SIDS2/gcgdata/geneseq/geneseqn/NA2001.DAT:AAC99704

seq_documentation_block:
ID   AAC99704 standard; cDNA; 1421 BP.
XX
AC   AAC99704;
XX
DT   08-MAR-2001 (first entry)
XX
DE   Skin cell cDNA, SEQ ID NO: 244.
XX
KW   Mouse; skin cell; cytostatic; antiinflammatory; anti-HIV;
KW   nootropic; neuroprotective; vulnerary; immunomodulatory; vaccine;
KW   keratinocyte growth stimulation; cancer; angiogenesis inhibition;
KW   inflammation; neurological disease; ss.
XX
OS   Mus sp.
XX
PN   WO200069884-A2.
XX
PD   23-NOV-2000.
XX
PF   15-MAY-2000; 2000WO-NZ00075.
XX
PR   14-MAY-1999;   99US-0312283.
XX
PA   (GENE-) GENESIS RES & DEV CORP LTD.
XX
PI   Watson JD,  Strachan L,  Onrust R,  Sleeman M,  Kumble KD,  Murison JG;
XX
DR   WPI; 2001-007495/01.
DR   P-PSDB; AAB55958.
XX
PT   New isolated polynucleotide used in the identification of genetic
PT   disorders and encoding polypeptides used for treating inflammatory
PT   disease, cancer and neurological diseases -
XX
PS   Claim 1; Page 206-207; 352pp; English.
XX
CC   The present polynucleotide encodes a polypeptide which is expressed in
```

```
CC   mammalian skin cells. The polypeptide is useful for stimulating
CC   keratinocyte growth and motility, inhibiting the growth of cancer cells,
CC   modulating angiogenesis, inhibiting angiogenesis and vascularisation of
CC   tumours, modulating skin inflammation, stimulating the growth of
CC   epithelial cells, inhibiting the binding of human immunodeficiency virus
CC   (HIV)-1 to leukocytes, and treating inflammmatory disease, cancer and
CC   neurological diseases. The polynucleotide can be used as a marker, in
CC   the identification of genetic disorders, and for the design of
CC   oligonucleotides for examining expression patterns.
XX
SQ   Sequence 1421 BP; 290 A; 435 C; 429 G; 265 T; 2 other;


alignment_scores:
           Quality:  112.50              Length:    209
             Ratio:  1.103                 Gaps:     13
 Percent Similarity:  48.804   Percent Identity: 28.230

alignment_block:
 US-09-508-832-6 x AAC99704   ..

 Align seg 1/1  to: AAC99704  from: 1  to: 1421

       3 LysGlnProSerAspValSerSerGluCysAspArgGlu......GlyGl 17
         |||      |||:::   :::::::|||   :::::|||      :::||
     738 AAGGGTGGGTCGGAGGAGGAGGAGGAGAAGGAAGAGGAGGTGTTCAGTGG 787

      17 yGlnLeuGlnProAlaGluArgPro............ProGlnLeuArgP 30
         ||||   :::|||   |||  |||                 :::   :::|
     788 CCAGGAGGAGCCCCGGGAAGCCCCAGTATGTGGGGTCACTGAAGAGAAGC 837

      30 roGlyAlaPro...ThrSerLeuGlnThrGluProGlnGlyAsnPro... 44
       *   ||   :::|||      :::   :::::::|||   :::|||   |||
     838 CGGAGGTCCCTGACGAGACAGCCTCAGCAGAGGCTGAAGGGGTTCCCGCA 887

      45 ......AspGlyGluGlyAspArgCysProHisGlySer........... 55
               :::|||:::|||:::      |||:::|||||
     888 GCCAGCGAGGGCCAAGGGGAA......CCAGAAGGGTCTTTCTCCTTAGC 931

      56 .......ProGlnGlyProLeuAlaPro.ProAlaSerProGlyProPhe 69
               |||||||||      :::|||  |||   :::|||   |||
     932 CCAGGAACCCCAGGGAGCAGCTGGTCCTTCCGAAAGGTCCTGTGCCT... 978

      70 AlaThrArgSerProLeu.............................. 75
         ||||||:::|||||||||
     979 GCAACAGAATCTCCCCTAATGTGTAACAGGCCCCAGAACTGTGAGGCCTG 1028

      76 .PheIlePheValArgArgSerSerLeuLeuSerArgSerSerSerGlyT 92
          ::::::      |||||||||          ||||||::::::
    1029 ACTCTTGGGTCCTCGAAGGTCACCTCCTTGGTCAAGAAAGGCA....... 1071

      92 yrPheSerPheAspThrAspArgSerPro..................... 101
           ||||||||||||          :::|||
    1072 ..TTCAGCTTTGACTGCTTCTTGACACCCTGCCTTGGCCATTGTGGGTGC 1119

     102 ..........................AlaProMetSerCysAspLysSe 109
                                    |||   :::   |||      ::
    1120 CAATCCTGACCCTGAATGGGCAAAGCTGCTGGCCTCTGGTGTACCCCAGG 1169

     109 rThrGlnThrProSerProProCysGlnAlaPheAsnHisTyrLeuSerA 126
         ::::   |||||||||      :::          :::|||   |||:::|
    1170 AAACACCACCCCAAGTTCCAGCGCCCTTAATGACTCTCACATCCTGGGGG 1219

     126 laMetAlaSerIleArg..............GlnSerGlnGluGluPro 137
         ||                          :::::::|||   :::|||
    1220 CTTCACCCCGAAGCACCACTTTTCTGGAAGGGGAAGGTCAGACACATCCC 1269

     138 GluAspLeuArgProGluIleArgIleAlaGlnGluLeuArgArgIleGl 154
            :::|||:::|||:::            |||:::   |||      ||
    1270 ...AGTTTGGAGCCGCAATGAGGCAGTCCTCAGAACAGAAGGGGAACAGG 1316

     154 yAspGluPheAsnGluThrTyrThr 162
         |::::::::::::   |||||||||
```

```
    1317 CCAGAGGCTGACTGTGACATACACA 1341

seq_name: /SIDS2/gcgdata/geneseq/geneseqn/NA2001.DAT:AAD07131

seq_documentation_block:
ID   AAD07131 standard; cDNA; 2800 BP.
XX
AC   AAD07131;
XX
DT   06-AUG-2001  (first entry)
XX
DE   Canine retinitis pigmentosa GTPase regulator (RPGR) mutant #1 cDNA.
XX
KW   Dog; X-linked progressive retinal atrophy 1; XLPRA1; genetic marker;
KW   retinitis pigmentosa GTPase regulator; RPGR; Siberian Husky; Samoyed;
KW   Miniature Schnauzer; mutant; mutein; ss.
XX
OS   Canis familiaris.
OS   Synthetic.
XX
FH   Key             Location/Qualifiers
FT   CDS             1..882
FT                   /*tag=  a
FT                   /product= "Canine retinitis pigmentosa GTPase
FT                   regulator mutant"
FT                   /note= "CDS does not include start codon"
FT                   /partial
FT   mutation        replace (877..878, AGAGAAT)
FT                   /*tag=  b
FT                   /note= "This deletion replaces Arg with Ile and results
FT                   in a premature stop codon"
XX
PN   WO200138578-A1.
XX
PD   31-MAY-2001.
XX
PF   21-NOV-2000; 2000WO-US31940.
XX
PR   24-NOV-1999;   99US-0167365.
XX
PA   (CORR ) CORNELL RES FOUND INC.
XX
PI   Aguirre GD,  Acland GM,  Zhang Q,  Ray K,  Zeiss CJ;
XX
DR   WPI; 2001-367707/38.
DR   P-PSDB; AAE02398.
XX
PT   Identifying dogs with or carrying X-linked progressive retinal atrophy
PT   by detecting retinitis pigmentosa GTPase regulator gene mutation,
PT   useful when breeding Husky, Samoyed and Miniature Schnauzer  -
XX
PS   Claim 73; Page 32-33; 88pp; English.
XX
CC   The invention relates to a method for identifying dogs which are
CC   genetically normal, are carriers of, or are affected with X-linked
CC   progressive retinal atrophy (XLPRA), by testing a biological sample with
CC   genetic markers that co-segregate with a XLPRA gene locus. The invention
CC   also relates to canine retinitis pigmentosa GTPase regulator (RPGR), RPGR
CC   mutants and their corresponding nucleic acid molecules. The mutated RPGR
CC   genes are responsible for the XLPRA in dogs. Methods are used to select
CC   dogs for breeding so that dogs carrying the mutated locus are eliminated
CC   from the breeding stock. The method particularly applies to Siberian
CC   Husky, Samoyed and Miniature Schnauzer breeds, or any other breed where
CC   the disease is X-linked. XLPRA1 type is identified in Siberian Huskies,
CC   and Samoyeds, while XLPRA2 type is identified in Miniature Schnauzers.
CC   The present cDNA sequence is the mutant open reading frame (ORF) 15
CC   encoding Canine retinitis pigmentosa GTPase regulator (RPGR) mutant
CC   found in XLPRA1-affected dogs. This mutant is obtained by deleting
CC   'GAGAA' nucleotide bases from position 878 to 882 of the wild-type
CC   canine RPGR cDNA.
XX
SQ   Sequence 2800 BP; 1026 A; 287 C; 938 G; 549 T; 0 other;
```

```
alignment_scores:
            Quality:  110.00              Length:    147
              Ratio:  1.279                 Gaps:      6
 Percent Similarity:  58.503    Percent Identity: 27.891

alignment_block:
 US-09-508-832-6 x AAD07131/rev   ..

 Align seg 1/1  to reverse of: AAD07131  from: 1  to: 2800

       5 ProSerAspValSerSerGluCysAspArgGluGlyGlyGlnLeuGlnPr 21
         ||||||::::::||||||:::   :::   :::   :::
     770 CCTTCTTCCACTTCCTCCTCTCCTTCCTCCTCCACTTCCCCTTCTCCTTC 721

      21 oAlaGluArgProProGlnLeuArgProGlyAlaProThrSerLeuGlnT 38
          :::     |||         :::   ::::::|||:::|||     :
     720 CTCCACTTCCCCCTTTCCTTCCCATTCTGCCTCTCCTTCCTCCCCTCCAG 671

      38 hrGluProGlnGlyAsnProAspGlyGlu...GlyAspArgCysProHis 53
         ::::::|||   ::::::|||:::::::::   ::::::      |||
     670 CCTCCCCTTTCTCCTCCCCTTCCTCTTCCCCCTCCTCCTCTTCCCCCTCC 621

      54 GlySerPro......GlnGlyProLeuAlaProProAla......SerPr 66
         :::||||||      :::::::|||:::      |||:::      |||||
     620 AATTCCCCCTCTCTTTCTTCCCCTTTTCCCTCTCCTTCCTCTAATTCCCC 571

      66 oGlyProPheAlaThrArgSerProLeuPheIlePheValArgArgSerS 83
         |:::|||   :::::::::||||||                        |
     570 CTCCCCTTCCTCTTCCAATTCCCCC.....................T 545

      83 erLeuLeuSerArgSerSerSerGlyTyrPheSerPheAspThrAspArg 99
         ||      |||   :::||||||:::   |||||||||::::::
     544 CTCCTTCCTCCTCTACTTCTTCTGCTTCCTTCTCTTTCTCTGCTCCTTCC 495

     100 Ser.....................ProAlaProMetSerCysAspLy 108
         |||                        |||:::|||         :::::
     494 TCTTCCCCTTCCCCTTCTCTTTTCTTCCCTTCCCCCTCTCCTCCTTCCTC 445

     108 sSerThrGlnThrPro...SerProProCysGlnAlaPheAsnHisTyrL 124
         :|||:::   :::|||   ||||||||||||:::|||   :::   :::
     444 CTCTTCCCCCTCTCCTTCCTCCCCTCCTTGCTCTGCTTCTTCTTCTTTCT 395

     124 euSerAlaMetAlaSerIleArgGlnSerGlnGluGluPro 137
           ||||||   :::|||      :::|||:::   :::|||
     394 CCTCTGCTTCCTCCTCCTCTCCTTCTTCTTCCCCCTCTCCC 354

seq_name: /SIDS2/gcgdata/geneseq/geneseqn/NA2001.DAT:AAD07132

seq_documentation_block:
ID   AAD07132 standard; cDNA; 2803 BP.
XX
AC   AAD07132;
XX
DT   06-AUG-2001  (first entry)
XX
DE   Canine retinitis pigmentosa GTPase regulator (RPGR) mutant #2 cDNA.
XX
KW   Dog; X-linked progressive retinal atrophy 2; XLPRA2; genetic marker;
KW   retinitis pigmentosa GTPase regulator; RPGR; Siberian Husky; Samoyed;
KW   Miniature Schnauzer; mutant; mutein; ss.
XX
OS   Canis familiaris.
OS   Synthetic.
XX
FH   Key             Location/Qualifiers
FT   CDS             1..1149
FT                   /*tag=  a
FT                   /product= "Canine retinitis pigmentosa GTPase
FT                   regulator mutant"
FT                   /note= "CDS does not include start codon"
FT                   /partial
FT   mutation        replace (931..932, AGAG)
FT                   /*tag=  b
FT                   /note= "This deletion results in the change of
FT                   amino acids and ends in a premature stop codon"
XX
PN   WO200138578-A1.
XX
PD   31-MAY-2001.
XX
PF   21-NOV-2000; 2000WO-US31940.
XX
PR   24-NOV-1999;  99US-0167365.
XX
PA   (CORR ) CORNELL RES FOUND INC.
XX
PI   Aguirre GD,  Acland GM,  Zhang Q,  Ray K,  Zeiss CJ;
XX
DR   WPI; 2001-367707/38.
DR   P-PSDB; AAE02399.
XX
PT   Identifying dogs with or carrying X-linked progressive retinal atrophy
PT   by detecting retinitis pigmentosa GTPase regulator gene mutation,
PT   useful when breeding Husky, Samoyed and Miniature Schnauzer
XX
PS   Claim 75; Page 34-35; 88pp; English.
XX
CC   The invention relates to a method for identifying dogs which are
CC   genetically normal, are carriers of, or are affected with X-linked
CC   progressive retinal atrophy (XLPRA), by testing a biological sample with
CC   genetic markers that co-segregate with a XLPRA gene locus. The invention
CC   also relates to canine retinitis pigmentosa GTPase regulator (RPGR), RPGR
CC   mutants and their corresponding nucleic acid molecules. The mutated RPGR
CC   genes are responsible for the XLPRA in dogs. Methods are used to select
CC   dogs for breeding so that dogs carrying the mutated locus are eliminated
CC   from the breeding stock. The method particularly applies to Siberian
CC   Husky, Samoyed and Miniature Schnauzer breeds, or any other breed where
CC   the disease is X-linked. XLPRA1 type is identified in Siberian Huskies,
CC   and Samoyeds, while XLPRA2 type is identified in Miniature Schnauzers.
CC   The present cDNA sequence is the mutant open reading frame (ORF) 15
CC   encoding Canine retinitis pigmentosa GTPase regulator (RPGR) mutant
CC   found in XLPRA2-affected dogs. This mutant is obtained by deleting
CC   'GA' nucleotide bases from position 932 to 933 of the wild-type
CC   canine RPGR cDNA.
XX
SQ   Sequence 2803 BP; 1028 A; 287 C; 939 G; 549 T; 0 other;


alignment_scores:
            Quality:  110.00              Length:    147
              Ratio:  1.279                 Gaps:      6
 Percent Similarity:  58.503    Percent Identity: 27.891

alignment_block:
 US-09-508-832-6 x AAD07132/rev   ..

 Align seg 1/1  to reverse of: AAD07132  from: 1  to: 2803

       5 ProSerAspValSerSerGluCysAspArgGluGlyGlyGlnLeuGlnPr 21
         ||||||::::::||||||:::   :::   :::   :::
     770 CCTTCTTCCACTTCCTCCTCTCCTTCCTCCTCCACTTCCCCTTCTCCTTC 721

      21 oAlaGluArgProProGlnLeuArgProGlyAlaProThrSerLeuGlnT 38
          :::     |||         :::   ::::::|||:::|||     :
     720 CTCCACTTCCCCCTTTCCTTCCCATTCTGCCTCTCCTTCCTCCCCTCCAG 671

      38 hrGluProGlnGlyAsnProAspGlyGlu...GlyAspArgCysProHis 53
         ::::::|||   ::::::|||:::::::::   ::::::      |||
     670 CCTCCCCTTTCTCCTCCCCTTCCTCTTCCCCCTCCTCCTCTTCCCCCTCC 621

      54 GlySerPro......GlnGlyProLeuAlaProProAla......SerPr 66
         :::||||||      :::::::|||:::      |||:::      |||||
     620 AATTCCCCCTCTCTTTCTTCCCCTTTTCCCTCTCCTTCCTCTAATTCCCC 571

      66 oGlyProPheAlaThrArgSerProLeuPheIlePheValArgArgSerS 83
```

```
      |:::|||   :::::::::||||||                          |
    570 CTCCCCTTCCTCTTCCAATTCCCCC.....................T 545

     83 erLeuLeuSerArgSerSerSerGlyTyrPheSerPheAspThrAspArg 99
        ||      |||   :::|||||||:::   |||||||||:::::::
    544 CTCCTTCCTCCTCTACTTCTTCTGCTTCCTTCTCTTTCTCTGCTCCTTCC 495

    100 Ser.....................ProAlaProMetSerCysAspLy 108
        |||                        |||:::|||         :::::
    494 TCTTCCCCTTCCCCTTCTCTTTTCTTCCCTTCCCCCTCTCCTCCTTCCTC 445

    108 sSerThrGlnThrPro...SerProProCysGlnAlaPheAsnHisTyrL 124
        :|||:::   :::|||   ||||||||||||:::|||   :::   :::
    444 CTCTTCCCCCTCTCCTTCCTCCCCTCCTTGCTCTGCTTCTTCTTCTTTCT 395

    124 euSerAlaMetAlaSerIleArgGlnSerGlnGluGluPro 137
          ||||||   :::|||      :::|||:::   :::|||
    394 CCTCTGCTTCCTCCTCCTCTCCTTCTTCTTCCCCCTCTCCC 354

seq_name: /SIDS2/gcgdata/geneseq/geneseqn/NA2001.DAT:AAD07130

seq_documentation_block:
ID   AAD07130 standard; cDNA; 2805 BP.
XX
AC   AAD07130;
XX
DT   06-AUG-2001  (first entry)
XX
DE   Canine retinitis pigmentosa GTPase regulator (RPGR) cDNA.
XX
KW   Dog; X-linked progressive retinal atrophy; XLPRA; genetic marker;
KW   retinitis pigmentosa GTPase regulator; RPGR; Siberian Husky; Samoyed;
KW   Miniature Schnauzer; ss.
XX
OS   Canis familiaris.
XX
FH   Key             Location/Qualifiers
FT   CDS             1..1569
FT                   /*tag=  a
FT                   /product= "Canine retinitis pigmentosa GTPase regulator"
FT                   /note= "CDS does not include start codon"
FT                   /partial
XX
PN   WO200138578-A1.
XX
PD   31-MAY-2001.
XX
PF   21-NOV-2000; 2000WO-US31940.
XX
PR   24-NOV-1999;   99US-0167365.
XX
PA   (CORR ) CORNELL RES FOUND INC.
XX
PI   Aguirre GD,  Acland GM,  Zhang Q,  Ray K,  Zeiss CJ;
XX
DR   WPI; 2001-367707/38.
DR   P-PSDB; AAE02397.
XX
PT   Identifying dogs with or carrying X-linked progressive retinal atrophy
PT   by detecting retinitis pigmentosa GTPase regulator gene mutation,
PT   useful when breeding Husky, Samoyed and Miniature Schnauzer  -
XX
PS   Claim 71; Page 30-31; 88pp; English.
XX
CC   The invention relates to a method for identifying dogs which are
CC   genetically normal, are carriers of, or are affected with X-linked
CC   progressive retinal atrophy (XLPRA), by testing a biological sample with
CC   genetic markers that co-segregate with a XLPRA gene locus. The invention
CC   also relates to canine retinitis pigmentosa GTPase regulator (RPGR), RPGR
CC   mutants and their corresponding nucleic acid molecules. The mutated RPGR
CC   genes are responsible for the XLPRA in dogs. Methods are used to select
CC   dogs for breeding so that dogs carrying the mutated locus are eliminated
CC   from the breeding stock. The method particularly applies to Siberian
CC   Husky, Samoyed and Miniature Schnauzer breeds, or any other breed where
CC   the disease is X-linked. XLPRA1 type is identified in Siberian Huskies,
CC   and Samoyeds, while XLPRA2 type is identified in Miniature Schnauzers.
CC   The present cDNA sequence is the normal open reading frame (ORF) 15
CC   encoding Canine retinitis pigmentosa GTPase regulator (RPGR) protein
CC   found in normal dogs.
XX
SQ   Sequence 2805 BP; 1029 A; 287 C; 940 G; 549 T; 0 other;

alignment_scores:
            Quality:  110.00              Length:    147
              Ratio:  1.279                 Gaps:      6
 Percent Similarity:  58.503   Percent Identity: 27.891

alignment_block:
 US-09-508-832-6 x AAD07130/rev   ..

 Align seg 1/1  to reverse of: AAD07130  from: 1  to: 2805

      5 ProSerAspValSerSerGluCysAspArgGluGlyGlyGlnLeuGlnPr 21
        ||||||::::::||||||:::   :::   :::   :::
    770 CCTTCTTCCACTTCCTCCTCTCCTTCCTCCTCCACTTCCCCTTCTCCTTC 721

     21 oAlaGluArgProProGlnLeuArgProGlyAlaProThrSerLeuGlnT 38
          :::     |||         :::   ::::::|||:::|||      :
    720 CTCCACTTCCCCCTTTCCTTCCCATTCTGCCTCTCCTTCCTCCCCTCCAG 671

     38 hrGluProGlnGlyAsnProAspGlyGlu...GlyAspArgCysProHis 53
        :::::|||   :::::::|||:::::::::   ::::::      |||
    670 CCTCCCCTTTCTCCTCCCCTTCCTCTTCCCCCTCCTCCTCTTCCCCCTCC 621

     54 GlySerPro......GlnGlyProLeuAlaProProAla......SerPr 66
        :::||||||      ::::::|||:::      |||:::      |||||
    620 AATTCCCCCTCTCTTTCTTCCCCTTTTCCCTCTCCTTCCTCTAATTCCCC 571

     66 oGlyProPheAlaThrArgSerProLeuPheIlePheValArgArgSerS 83
        |:::|||   :::::::::||||||                        |
    570 CTCCCCTTCCTCTTCCAATTCCCCC.....................T 545

     83 erLeuLeuSerArgSerSerSerGlyTyrPheSerPheAspThrAspArg 99
        ||      |||   :::|||||||:::   |||||||||:::::::
    544 CTCCTTCCTCCTCTACTTCTTCTGCTTCCTTCTCTTTCTCTGCTCCTTCC 495

    100 Ser.....................ProAlaProMetSerCysAspLy 108
        |||                        |||:::|||         :::::
    494 TCTTCCCCTTCCCCTTCTCTTTTCTTCCCTTCCCCCTCTCCTCCTTCCTC 445

    108 sSerThrGlnThrPro...SerProProCysGlnAlaPheAsnHisTyrL 124
        :|||:::   :::|||   ||||||||||||:::|||   :::   :::
    444 CTCTTCCCCCTCTCCTTCCTCCCCTCCTTGCTCTGCTTCTTCTTCTTTCT 395

    124 euSerAlaMetAlaSerIleArgGlnSerGlnGluGluPro 137
          ||||||   :::|||      :::|||:::   :::|||
    394 CCTCTGCTTCCTCCTCCTCTCCTTCTTCTTCCCCCTCTCCC 354

seq_name: /SIDS2/gcgdata/geneseq/geneseqn/NA1996.DAT:AAT18023

seq_documentation_block:
ID   AAT18023 standard; cDNA; 6171 BP.
XX
AC   AAT18023;
XX
DT   19-MAY-1996  (first entry)
XX
DE   Rat RIZ cDNA.
XX
KW   Retinoblastoma protein-interacting zinc finger; RIZ;
KW   cell proliferation; tumour; cancer; neuroblastoma; melanoma;
KW   diagnosis; therapy; ss.
XX
OS   Rattus sp.
XX
```

```
FH   Key             Location/Qualifiers
FT   CDS             157..5277
FT                   /*tag=  a
XX
PN   WO9606168-A2.
XX
PD   29-FEB-1996.
XX
PF   18-AUG-1995;    95WO-US10574.
XX
PR   06-MAR-1995;    95US-0399411.
PR   18-AUG-1994;    94US-0292683.
XX
PA   (LJOL-) LA JOLLA CANCER RES FOUND.
XX
PI   Huang S;
XX
DR   WPI; 1996-151371/15.
DR   P-PSDB; AAR92103.
XX
PT   Nucleic acid encoding mammalian retinoblastoma protein-interacting
PT   zinc finger  - used to identify cell proliferation modulating agents
PT   for treatment of tumours, esp. neuroblastoma and melanoma, also used
PT   for cancer diagnosis.
XX
PS   Claim 10; Fig 1A; 142pp; English.
XX
CC   A complete rat retinoblastoma (Rb) protein-interacting zinc finger
CC   (RIZ) cDNA sequence (AAT18023) was cloned from a rat neonatal cardiac
CC   myocyte lambda-gt11 cDNA expression library.  At least 2 forms of
CC   RIZ mRNA are expressed in rat brain, suggesting that RIZ mRNA can
CC   arise by alternative splicing.  Rat RIZ-encoding sequences are used
CC   for prodn. of rat RIZ (AAR92103) or its active fragments by recombinant
CC   DNA methods, and in the diagnosis of disorders associated with
CC   abnormal levels of RIZ expression.
XX
SQ   Sequence 6171 BP; 1683 A; 1645 C; 1524 G; 1319 T; 0 other;


alignment_scores:
            Quality:  109.50              Length:    222
              Ratio:  1.084                 Gaps:     12
 Percent Similarity:  45.495   Percent Identity: 24.775

alignment_block:
 US-09-508-832-6 x AAT18023    ..

 Align seq 1/1  to: AAT18023  from: 1  to: 6171

       2 AlaLysGlnProSerAspVal......SerSerGluCysAspArgGluGl 16
         |||::::::|||   :::|||         |||:::         :::
    2941 GCTGAGTCTCCACCTGAAGTTGTTGGCCCTTCCTCACCCCCTCTCCAGAC 2990

      16 yGlyGlnLeuGlnProAlaGluArgProProGlnLeuArgProGlyAlaP 33
         ::::::|||:::   ::::::   ||||||   |||   |||      |
    2991 AGCCTCCTTATCCTCCGGTCAGCTGCCTCCTCTCTTAACCCCCACAGAGC 3040

      33 roThrSer..................................... 35
         ||:::|||
    3041 CTTCTTCCCCTCCCCCCTGCCCTCCTGTGTTAACTGTTGCCACTCCACCA 3090

      36 .........LeuGlnThrGluProGlnGlyAsnPro......AspGlyGl 47
                 |||   |||   |||   :::::::|||      |||:::::
    3091 CCTCCCCTCCTTCCAACCGTCCCTCTCTCCCACCCCTCTTCTGATGCCTC 3140

      47 uGlyAspArgCysProHis.....................GlySerProG 57
         :   ::::::||||||                        ||||||
    3141 CCCTCAGCAGTGTCCCTCTCCGTTCTCAAACACCACTGCTCAGTCTCCTC 3190

      57 lnGlyProLeuAlaProProAlaSerProGlyPro.............. 68
               |||:::|||   :::||||||:::|||
    3191 TTCCCATTCTCTCCCCAACAGTGTCTCCCTCTCCCTCTCCCATTCCTCCT 3240

      69 ............PheAlaThrArgSerProLeuPheIlePheValArgAr 81
                     :::::::::   ||||||
    3241 GTAGAGCCACTTATGTCTGCTGCTTCCCCT.................GG 3272

      81 gSerSerLeuLeuSerArgSerSerSerGlyTyrPheSerPheAspThrA 98
                |||||   |||||||||:::      ||||||   ::::
    3273 TCCCCCAACACTTTCTTCCTCCTCCTCTTCTTCCTCTTCCTTCCCTTCCT 3322

      98 spArgSerProAlaProMetSerCysAspLysSerThrGlnThrProSer 114
         ::                |||||      |||:::   :::||||||
    3323 CT..................TCCTGC......TCCTCCACCTCCCCCTCC 3348

     115 ProProCysGlnAlaPheAsnHisTyrLeuSerAlaMetAlaSerIle.. 130
         ||||||                    |||||||||:::::|||:::
    3349 CCACCC.................CCTCTTTCAGCAGTGTCATCTGTGGT 3380

     131 ..............................................ArgG 132
                                                      :::|
    3381 TTCCTCTGGGGACAACCTGGAGGCATCTCTGCCTGCAGTAACTTTCAAAC 3430

     132 lnSerGlnGluGluProGluAspLeuArgProGluIleArgIleAlaGln 148
         ||::::::::::|||   |||   |||:::|||:::          :::
    3431 AGGAGGAGTCAGAGAGTGAAGGTCTGAAACCCAAG............GAA 3468

     149 GluLeuArgArgIleGlyAspGlu......PheAsnGluThrTyrThrAr 163
         |||            |||   :::         :::|||||||:::::::
    3469 GAGGCCCCACCTGCAGGGGGACAGAGTGTGGTCCAAGAAACATTCAGCAA 3518

     163 gArgValPheAlaAsn 168
         ::::   ::::::|||
    3519 AAACTTCATTTGCAAT 3534

seq_name: /SIDS2/gcgdata/geneseq/geneseqn/NA2000.DAT:AAA60103

seq_documentation_block:
ID   AAA60103 standard; DNA; 6171 BP.
XX
AC   AAA60103;
XX
DT   17-JAN-2001  (first entry)
XX
DE   Rat Rb-interacting zinc finger (RIZ) protein coding sequence.
XX
KW   Rat; Rb-interacting zinc finger; RIZ; retinoblastoma; heart disease;
KW   cell proliferation; cell differentiation; tissue repair;
KW   transcription regulator; breast cancer; gene therapy; melanoma;
KW   neuroblastoma; leukaemia; Parkinson's disease; Huntingdon's disease;
KW   Alzheimer's disease; paralysis; motor neurone disorder; ds.
XX
OS   Rattus sp.
XX
FH   Key             Location/Qualifiers
FT   misc_feature    100..102
FT                   /*tag=  a
FT                   /note= "Upstream in-frame stop codon"
FT   CDS             157..5277
FT                   /*tag=  b
FT                   /product= "Rat RIZ protein"
XX
PN   US6069231-A.
XX
PD   30-MAY-2000.
XX
PF   18-AUG-1995;    95US-0516859.
XX
PR   18-AUG-1994;    94US-0292683.
PR   06-MAR-1995;    95US-0399411.
XX
PA   (LJOL-) LA JOLLA CANCER RES FOUND.
XX
PI   Huang S;
XX
DR   WPI; 2000-410879/35.
```

```
DR   P-PSDB; AAB12028.
XX
PT   New PR domain peptides comprising amino acid sequences from, for
PT   example retinoblastoma-interacting zinc finger, or egl-43 proteins, for
PT   regulating gene transcription and controlling cell proliferation and
PT   differentiation -
XX
PS   Example 1; Fig 1; 91pp; English.
XX
CC   The present sequence is the rat retinoblastoma (Rb)-interacting zinc
CC   finger (RIZ) protein coding sequence. RIZ is a nuclear phosphoprotein
CC   that acts as a cell differentiation factor. RIZ can modulate cell growth
CC   by binding to Rb protein, which is involved in regulating cell
CC   proliferation. In addition, RIZ can act to regulate transcription. RIZ
CC   functions to maintain cells in the G1 phase of the cell cycle, by
CC   interacting with Rb through the cr2 domain of RIZ. Rat RIZ protein
CC   contains a number of GTPase motifs (see AAB12037 to AAB12056 and AAB12099
CC   to AAB12104). RIZ protein is a PR domain protein and is present primarily
CC   in the cell nucleus. RIZ gene mutations may be implicated in various
CC   cancers such as melanoma, neuroblastoma, leukaemia and breast cancer, and
CC   so the RIZ gene may be used in gene therapy for these disorders. Since
CC   RIZ protein is implicated in cell cycle arrest, inhibition of RIZ
CC   activity may be useful in neurodegenerative disorder therapy e.g. for
CC   Parkinson's, Huntingdon's or Alzheimer's disease, paralysis or motor
CC   neurone disorders, or cardiac disorder therapy e.g for heart disease
CC   where the ability to induce neural/ cardiac tissue proliferation would be
CC   useful.
XX
SQ   Sequence 6171 BP; 1683 A; 1645 C; 1524 G; 1319 T; 0 other;


alignment_scores:
      Quality:  109.50          Length:    222
        Ratio:  1.084             Gaps:     12
 Percent Similarity:  45.495   Percent Identity: 24.775

alignment_block:
 US-09-508-832-6 x AAA60103   ..

 Align seg 1/1  to: AAA60103  from: 1  to: 6171

       2 AlaLysGlnProSerAspVal......SerSerGluCysAspArgGluGl 16
         |||:::::::|||   :::|||         |||:::         :::
    2941 GCTGAGTCTCCACCTGAAGTTGTTGGCCCTTCCTCACCCCCTCTCCAGAC 2990

      16 yGlyGlnLeuGlnProAlaGluArgProProGlnLeuArgProGlyAlaP 33
          ::::::|||:::   ::::::   ||||||   |||   |||      |
    2991 AGCCTCCTTATCCTCCGGTCAGCTGCCTCCTCTCTTAACCCCCACAGAGC 3040

      33 roThrSer..................................... 35
         ||:::|||
    3041 CTTCTTCCCCTCCCCCCTGCCCTCCTGTGTTAACTGTTGCCACTCCACCA 3090

      36 .........LeuGlnThrGluProGlnGlyAsnPro......AspGlyGl 47
                 |||   |||   |||   :::::::|||      |||:::::
    3091 CCTCCCCTCCTTCCAACCGTCCCTCTCTCCCACCCCTCTTCTGATGCCTC 3140

      47 uGlyAspArgCysProHis.....................GlySerProG 57
         :   ::::::||||||                        ||||||
    3141 CCCTCAGCAGTGTCCCTCTCCGTTCTCAAACACCACTGCTCAGTCTCCTC 3190

      57 lnGlyProLeuAlaProProAlaSerProGlyPro.............. 68
                |||:::|||   :::|||||||:::|||
    3191 TTCCCATTCTCTCCCCAACAGTGTCTCCCTCTCCCTCTCCCATTCCTCCT 3240

      69 ............PheAlaThrArgSerProLeuPheIlePheValArgAr 81
                    :::::::::   ||||||
    3241 GTAGAGCCACTTATGTCTGCTGCTTCCCCT.................GG 3272

      81 gSerSerLeuLeuSerArgSerSerSerGlyTyrPheSerPheAspThrA 98
                 ||||||   |||||||||:::      ||||||   ::::
    3273 TCCCCCAACACTTTCTTCCTCCTCCTCTTCTTCCTCTTCCTTCCCTTCCT 3322

      98 spArgSerProAlaProMetSerCysAspLysSerThrGlnThrProSer 114
         ::                  ||||||      |||:::   :::||||||
    3323 CT.................TCCTGC......TCCTCCACCTCCCCCTCC 3348

     115 ProProCysGlnAlaPheAsnHisTyrLeuSerAlaMetAlaSerIle.. 130
         ||||||                    |||||||||:::::|||:::
    3349 CCACCC.................CCTCTTTCAGCAGTGTCATCTGTGGT 3380

     131 .....................................ArgG 132
                                              :::|
    3381 TTCCTCTGGGGACAACCTGGAGGCATCTCTGCCTGCAGTAACTTTCAAAC 3430

     132 lnSerGlnGluGluProGluAspLeuArgProGluIleArgIleAlaGln 148
         ||::::::::::|||   |||   |||:::|||:::            :::
    3431 AGGAGGAGTCAGAGAGTGAAGGTCTGAAACCCAAG............GAA 3468

     149 GluLeuArgArgIleGlyAspGlu......PheAsnGluThrTyrThrAr 163
         |||           |||   :::        :::|||||||::::::::
    3469 GAGGCCCCACCTGCAGGGGGACAGAGTGTGGTCCAAGAAACATTCAGCAA 3518

     163 gArgValPheAlaAsn 168
         ::::   ::::::|||
    3519 AAACTTCATTTGCAAT 3534

seq_name: /SIDS2/gcgdata/geneseq/geneseqn/NA2000.DAT:AAZ50085

seq_documentation_block:
ID   AAZ50085 standard; cDNA; 1249 BP.
XX
AC   AAZ50085;
XX
DT   04-MAY-2000  (first entry)
XX
DE   Rice serine O-acetyltransferase-1 cDNA clone.
XX
KW   Serine O-acetyltransferase; sulphate assimilation; O-acetylserine;
KW   rice; clone rlr24.pk0069.a11; cysteine formation; marker; probe;
KW   plant breeding; transgenic plant; ss.
XX
OS   Oryza sativa.
XX
FH   Key             Location/Qualifiers
FT   CDS             119..1030
FT                   /*tag= a
FT                   /product= "Rice serine O-acetyltransferase-1"
FT                   /note= "Derived from the clone rlr24.pk0069.a11"
FT                   /transl_except= (pos:194..196, aa:Val)
XX
PN   WO200004167-A2.
XX
PD   27-JAN-2000.
XX
PF   13-JUL-1999;   99WO-US15872.
XX
PR   14-JUL-1998;   98US-0092833.
XX
PA   (DUPO ) DU PONT DE NEMOURS & CO E I.
XX
PI   Falco SC,  Allen SM,  Maxwell CA;
XX
DR   WPI; 2000-182432/16.
DR   P-PSDB; AAY44768.
XX
PT   New isolated nucleic acid fragment encoding a sulfate assimilation
PT   protein in plants, useful as probes to isolate genes encoding
PT   homologous proteins from other plant species -
XX
PS   Claim 3; Page 35; 44pp; English.
XX
CC   The present sequence is a cDNA encoding rice serine O-acetyltransferase,
CC   a sulphate assimilation protein. This is obtained from rlr24 cDNA
CC   library, clone rlr24.pk0069.a11, derived from rice leaf infected with
CC   Magaporthe grisea strain 4360-R-62. Serine O-acetyltransferase converts
```

```
CC   serine to O-acetylserine, that is involved in the formation of cysteine.
CC   This sequence is used as a probe to isolate other plant sulphate
CC   assimilation proteins, for genetic and physical mapping of related genes
CC   and as markers of traits linked to the gene. This is useful for plant
CC   breeding. It is also used to create transgenic plants with altered
CC   levels of serine O-acetyltransferase, or found in cell types or
CC   developmental stages in which they are not normally found.
XX
SQ   Sequence 1249 BP; 193 A; 477 C; 361 G; 218 T; 0 other;


alignment_scores:
           Quality:  109.00            Length:    191
             Ratio:  1.239               Gaps:      8
 Percent Similarity:  46.073   Percent Identity: 27.225

alignment_block:
 US-09-508-832-6 x AAZ50085   ..

 Align seg 1/1  to: AAZ50085  from: 1  to: 1249

      25 ProProGlnLeuArgPro.................GlyAlaProThrSe 35
         ||||||:::      |||                  ::::::|||   ||
     378 CCTCCACGCTGCTCTCCACGCTCCTCTACGACCTCTTCGTCGCCTCCCTC 427

      35 rLeuGlnThrGluProGlnGlyAsnProAspGlyGluGlyAspArgCysP 52
         |   :::|||   |||::::::   |||::::::              |
     428 GCCGCGCACCCCACCCTCCGCGCCGCCGTCGTCG..............C 462

      52 roHisGlySerProGlnGlyProLeuAlaProProAlaSerProGlyPro 68
         ||   :::|||||||::::::|||   :::||||||||||||   :::|||
     463 CGACCTCCTCGCCGCGCGCTCCAGGGACCCCGCCTGCGTCGGCTTCTCCC 512

      69 PheAlaThrArgSerProLeuPheIlePheValArgArgSerSer..... 83
            |||:::   :::                :::|||   ||||||
     513 ACTGCCTCCTCAACT..............ACAAGGGCTTCCTCGCCATC 547

      84 .......LeuLeuSerArgSerSerSerGlyTyrPheSerPheAspThrA 98
                |||||||:::   ||||||      :::        :
     548 CAGGCCCAGCGCGTCGCGCACGTGCTCTGGGCGCAGGACCGCCGCGCCCT 597

      98 spArg.........SerProAlaProMetSerCysAspLysSerThrGln 111
         ::|||         |||||||||         |||:::   ||||||:::
     598 CGCGCTCGCGCTCCAGTCCCGCGTCGCCGAGGTGTTCGCCGTCGACATCC 647

     112 ThrProSerProProCysGlnAlaPheAsnHisTyrLeuSerAlaMetAl 128
         ||||||   |||                :::        :::      ||
     648 ACCCCGCCGCCGCGATCGGCAAGGGCGTCCTCCTCGACCACGCCACGGGC 697

     128 aSerIle.ArgGlnSerGlnGluGlu..................... 136
         ||||    |||:::::::::::::::
     698 GTCGTCATCGGAGAGACCGCCGTCATCGGCGACAACGTCTCCATCCTCCA 747

     137 .....................................ProGl 138
                                              |||::
     748 CCACGTCACGCTGGGCGGGACAGGCAAGGCCGTGGGCGACCGGCACCCCA 797

     138 uAsp......LeuArgProGluIleArgIleAlaGlnGluLeuArgArgI 153
         :|||         ||||||:::   |||      ::::::   |||:::
     798 AGATCGGCGACGGCGTCCTCATTGGCGCCGGCGCGACGATCCTCGGCAAT 847

     153 leGlyAspGluPheAsnGluThrTyrThrArgArgValPheAlaAsnAsp 169
            |||:::   :::              ||||||   :::
     848 GTCAGGATCGGCGCCGGGGCCAAGATCGGGGCCGGGTCGCTGGTGCTCAT 897

     170 TyrArgGluAlaGluAspHis 176
            |||   ||||||||||||
     898 CGACGTGCCGCCGAGGACCAC 918
```

```
OM of: US-09-508-832-6 to: EST:*   out_format : pfs

Date: Dec 11, 2001  1:03 AM

About: Results were produced by the GenCore software, version 4.5,
       Copyright (c) 1993-2000 Compugen Ltd.

Command line parameters:
 -MODEL=frame+_p2n.model -DEV=xlp
 -Q=/cgn2_1/USPTO_spool/US09508832/runat_10122001_110349_29536/app_query.fasta_1.620
 -DB=EST -QFMT=fastap -SUFFIX=rst -GAPOP=12.000 -GAPEXT=4.000
 -MINMATCH=0.100 -LOOPCL=0.000 -LOOPEXT=0.000 -QGAPOP=4.500
 -QGAPEXT=0.050 -XGAPOP=10.000 -XGAPEXT=0.500 -FGAPOP=6.000
 -FGAPEXT=7.000 -YGAPOP=10.000 -YGAPEXT=0.500 -DELOP=6.000
 -DELEXT=7.000 -START=1 -MATRIX=blosum62 -TRANS=human40.cdi
 -LIST=45 -DOCALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0
 -ALIGN=15 -MODE=LOCAL -OUTFMT=pfs -NORM=ext -MINLEN=0
 -MAXLEN=2000000000 -USER=US09508832_@CGN1_1_0 -NCPU=6 -ICPU=3
 -LONGLOG -NO_XLPXY -WAIT -THREADS=1

Search information block:
Query: US-09-508-832-6
Query length: 196
Database: EST:*
Database sequences: 11351937
Database length: 1077921985
Search time (sec): 2629.110000



score_list:
Sequence          Strd Orig    ZScore    EScore Len  ! Documentation  ..
gb_htc:AK011490     +  980.50 1136.71  3.8e-54   1206  ! AK011490 Mus musculus 10 days
gb_est2:BG921698    +  918.00 1066.42  3.1e-50    935  ! BG921698 602825518F1 NCI_CGAP_M
gb_est2:BF021882    +  730.00  854.67  1.9e-38    452  ! BF021882 uy59b09.y1 McCarrey Ed
gb_gss:AZ706148     +  689.00  805.25  1.1e-35    580  ! AZ706148 RPCI-23-227P3.TV RPCI-
gb_est2:BG173095    +  606.50  708.68  2.6e-30    668  ! BG173095 602336666F1 NCI_CGAP_M
gb_est1:AI971169    -  332.00  393.51  9.4e-13    492  ! AI971169 wr24h12.x1 NCI_CGAP_Pr
gb_est2:BF319454    -  249.00  299.35  1.6e-07    389  ! BF319454 uy59b09.x1 McCarrey Ed
gb_est1:AW629314    -  249.00  295.11  2.8e-07    664  ! AW629314 hi56e02.x1 Soares_NFL_
gb_est1:AF209718    +  202.00  242.41  0.0002     537  ! AF209718 AF209718 Xenopus laevi
gb_est1:AW748960    +  149.00  190.84  0.1822     157  ! AW748960 RC4-BT0312-081199-011-
gb_est1:AA629308    -  127.00  156.19  15.51      501  ! AA629308 zu84g06.s1 Soares_test
gb_est2:BF172831    +  126.00  161.92  7.43       210  ! BF172831 PCL5805 Myeloma (PCL)
gb_est1:AA629050    -  123.00  151.56  28.09      501  ! AA629050 zu84a06.s1 Soares_test
gb_est2:BF259468    +  118.50  142.97  84.57      768  ! BF259468 HVSMEf0019D03f Hordeum
gb_htc:BC007683     +  115.00  134.44  252.28    1350  ! BC007683 Homo sapiens, postmei
gb_est2:BF631591    -  114.50  137.32  174.41     873  ! BF631591 HVSMEb0016I15f Hordeum
gb_est1:BE408300    +  113.00  138.19  156.12     629  ! BE408300 601302689F1 NIH_MGC_21
gb_est2:BF627834    -  113.00  136.36  197.35     792  ! BF627834 HVSMEb0005O20f Hordeum
gb_est1:AW550178    +  111.50  137.99  160.09     518  ! AW550178 L0061B04-3 NIA Mouse E
gb_est1:AV437695    +  111.50  137.85  162.92     527  ! AV437695 AV437695 Porphyra yezo
gb_gss:AQ744259     -  111.50  133.80  274.11     879  ! AQ744259 HS_5508_A2_E11_T7A RPC
gb_gss:CNS037HK     +  110.50  132.71  315.02     871  ! AL231329 Tetraodon nigroviridis
gb_est2:BF868252    +  110.00  126.80  672.16    1706  ! BF868252 963107F07.x1 C. reinh
gb_est2:BG997598    -  109.50  136.84  185.42     447  ! BG997598 PM0-HT0911-210301-018-
gb_est1:AV611648    +  109.50  135.37  223.87     538  ! AV611648 AV611648 Bos taurus lu
gb_est2:BG921067    -  109.50  133.29  292.58     700  ! BG921067 602825319F1 NCI_CGAP_M
gb_est2:BG394426    +  109.50  127.29  631.68    1492  ! BG394426 602456946F1 NIH_MGC_1
gb_est2:BI289133    -  108.50  134.17  261.15     541  ! BI289133 UI-R-DK0-cfe-h-03-0-UI
gb_est2:BF491700    +  108.50  133.04  301.95     624  ! BF491700 AT28505.5prime AT Dros
gb_est2:BF312248    +  108.50  128.16  564.76    1155  ! BF312248 601898759F1 NIH_MGC_1
gb_est1:AI368415    -  108.00  130.91  396.88     759  ! AI368415 qy08b11.x1 NCI_CGAP_Br
gb_est2:BG067996    -  108.00  127.36  625.94    1188  ! BG067996 H3060F01-3 NIA Mouse
gb_est2:BE880795    +  108.00  127.29  631.30    1198  ! BE880795 601493220F1 NIH_MGC_6
gb_est2:BG895004    +  107.50  132.64  317.74     567  ! BG895004 355805 MARC 1PIG Sus s
gb_est1:BE313313    -  107.50  129.40  481.91     854  ! BE313313 601147639F1 NIH_MGC_19
gb_gss:CNS03IUC     +  107.50  129.31  487.07     863  ! AL246045 Tetraodon nigroviridis
gb_est1:BE619071    +  107.50  129.29  488.79     866  ! BE619071 601472685F1 NIH_MGC_68
gb_est2:BG827242    +  107.50  128.04  573.86    1014  ! BG827242 602749384F1 NIH_MGC_1
gb_est1:AU124084    -  107.00  131.97  346.51     574  ! AU124084 AU124084 NT2RM2 Homo s
gb_est1:AJ006538    +  107.00  130.86  399.37     660  ! AJ006538 AJ006538 Zea mays earl
gb_est1:AL508424    +  107.00  130.52  417.22     689  ! AL508424 AL508424 Hordeum vulga
gb_est2:BF496767    +  106.50  130.79  402.97     619  ! BF496767 AT10677.5prime AT D
gb_est2:BF504432    +  106.50  130.70  407.60     626  ! BF504432 AT05788.5prime AT D
gb_est2:BF498774    +  106.50  130.51  417.54     641  ! BF498774 AT13245.5prime AT D
gb_est2:BF506369    +  106.50  129.79  457.98     702  ! BF506369 AT08819.5prime AT D

seq_name: gb_htc:AK011490

seq_documentation_block:
LOCUS       AK011490     1206 bp    mRNA            HTC       05-JUL-2001
DEFINITION  Mus musculus 10 days embryo cDNA, RIKEN full-length enriched
            library, clone:2610020M23, full insert sequence.
ACCESSION   AK011490
VERSION     AK011490.1  GI:12847647
KEYWORDS    CAP trapper.
SOURCE      Mus musculus (strain:C57BL/6J) 10 days embryo cDNA to mRNA,
            clone_lib:RIKEN full-length enriched mouse cDNA library
            clone:2610020M23.
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 1206)
  AUTHORS   Carninci,P. and Hayashizaki,Y.
  TITLE     High-efficiency full-length cDNA cloning
  JOURNAL   Methods in enzymology. 303, 19-44 (1999)
  MEDLINE   99279253
   PUBMED   10349636
REFERENCE   2  (bases 1 to 1206)
  AUTHORS   Carninci,P., Shibata,Y., Hayatsu,N., Sugahara,Y., Shibata,K.,
            Itoh,M., Konno,H., Okazaki,Y., Muramatsu,M. and Hayashizaki,Y.
  TITLE     Normalization and subtraction of cap-trapper-selected cDNAs to
            prepare full-length cDNA libraries for rapid discovery of new genes
  JOURNAL   Genome research. 10 (10), 1617-1630 (2000)
  MEDLINE   20499374
   PUBMED   11042159
REFERENCE   3  (bases 1 to 1206)
  AUTHORS   Shibata,K., Itoh,M., Aizawa,K., Nagaoka,S., Sasaki,N., Carninci,P.,
            Konno,H., Akiyama,J., Nishi,K., Kitsunai,T., Tashiro,H., Itoh,M.,
            Sumi,N., Ishii,Y., Nakamura,S., Hazama,M., Nishine,T., Harada,A.,
            Yamamoto,R., Matsumoto,H., Sakaguchi,S., Ikegami,T., Kashiwagi,K.,
            Fujiwake,S., Inoue,K., Togawa,Y., Izawa,M., Ohara,E., Watahiki,M.,
            Yoneda,Y., Ishikawa,T., Ozawa,K., Tanaka,T., Matsuura,S., Kawai,J.,
            Okazaki,Y., Muramatsu,M., Inoue,Y., Kira,A. and Hayashizaki,Y.
  TITLE     RIKEN integrated sequence analysis (RISA) system--384-format
            sequencing pipeline with 384 multicapillary sequencer
  JOURNAL   Genome research. 10 (11), 1757-1771 (2000)
  MEDLINE   20530913
   PUBMED   11076861
REFERENCE   4  (bases 1 to 1206)
  AUTHORS   The RIKEN Genome Exploration Research Group Phase II Team and the
            FANTOM Consortium.
  TITLE     Functional annotation of a full-length mouse cDNA collection
  JOURNAL   Nature 409, 685-690 (2001)
REFERENCE   5  (bases 1 to 1206)
  AUTHORS   Adachi,J., Aizawa,K., Akahira,S., Akimura,T., Aono,H., Arai,A.,
            Arakawa,T., Carninci,P., Fukuda,S., Fukunishi,Y., Furuno,M.,
            Hanagaki,T., Hara,A., Hayatsu,N., Hiramoto,K., Hiraoka,T., Hori,F.,
            Imotani,K., Ishii,Y., Itoh,M., Izawa,M., Kato,H., Kawai,J.,
            Kojima,Y., Konno,H., Kouda,M., Koya,S., Kurihara,C., Matsuyama,T.,
            Miyazaki,A., Nishi,K., Nomura,K., Numazaki,R., Ohno,M., Okazaki,Y.,
            Okido,T., Owa,C., Saito,H., Saito,R., Sakai,C., Sakai,K., Sano,H.,
            Sasaki,D., Shibata,K., Shibata,Y., Shinagawa,A., Shiraki,T.,
            Sogabe,Y., Suzuki,H., Tagami,M., Tagawa,A., Takahashi,F.,
            Tanaka,T., Tejima,Y., Toya,T., Yamamura,T., Yasunishi,A.,
            Yoshida,K., Yoshino,M., Muramatsu,M. and Hayashizaki,Y.
  TITLE     Direct Submission
  JOURNAL   Submitted (10-JUL-2000) Yoshihide Hayashizaki, The Institute of
            Physical and Chemical Research (RIKEN), Laboratory for Genome
            Exploration Research Group, RIKEN Gemomic Sciences Center (GSC),
            RIKEN Yokohama Institute; 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama,
            Kanagawa 230-0045, Japan (E-mail:genome-res@gsc.riken.go.jp,
            URL:http://genome.gsc.riken.go.jp/, Tel:81-45-503-9222,
            Fax:81-45-503-9216)
COMMENT     Please visit our web site (http://genome.gsc.riken.go.jp/) for
```

```
          further details.
          cDNA library was prepared and sequenced in Mouse Genome
          Encyclopedia Project of Genome Exploration Research Group in Riken
          Genomic Sciences Center and Genome Science Laboratory in RIKEN.
          Division of Experimental Animal Research in Riken contributed to
          prepare mouse tissues. First strand cDNA was primed with a primer
          [5' GAGAGAGAGAAGGATCCAAGAGCTCTTTTTTTTTTTTTTTVN 3'], cDNA was
          prepared by using trehalose thermo-activated reverse transcriptase
          and subsequently enriched for full-length by cap-trapper. cDNA went
          through one round of normalization to Rot = 7.5 and subtraction to
          Rot = 37.5. Second strand cDNA was prepared with the primer adapter
          of sequence [5'
          GAGAGAGAGATTCTCGAGTTAATTAAATTAATCCCCCCCCCCCCC 3']. cDNA was cleaved
          with XhoI and SstI. Cloning sites, 5' end: XhoI; 3' end: SstI.
          Host: SOLR.
FEATURES             Location/Qualifiers
     source          1. .1206
                     /organism="Mus musculus"
                     /strain="C57BL/6J"
                     /db_xref="taxon:10090"
                     /db_xref="MGD:MGI:1197519"
                     /db_xref="MGD:MGI:1902115"
                     /clone="2610020M23"
                     /clone_lib="RIKEN full-length enriched mouse cDNA library"
                     /dev_stage="10 days embryo"
BASE COUNT      265 a    339 c    298 g    301 t      3 others
ORIGIN


alignment_scores:
           Quality:  980.50            Length:    197
             Ratio:  5.161               Gaps:      2
 Percent Similarity:  96.447   Percent Identity: 95.939

alignment_block:
 US-09-508-832-6 x AK011490   ..

 Align_seg 1/1  to: AK011490  from: 1  to: 1206

       1 MetAlaLysGlnProSerAspValSerSerGluCysAspArgGluGlyGl 17
         ||||||||||||||||||||||||   ||||||||||||   ||||||||
     220 ATGGCCAAGCAACCTTCTGATGTACGTTCTGAGTGTGACAGCGAAGGTGG 269

      17 yGlnLeuGlnProAlaGluArgProProGlnLeuArgProGlyAlaProT 34
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     270 ACAATTGCAGCCTGCTGAGAGGCCTCCCCAGCTCAGGCCTGGGGCCCCTA 319

      34 hrSerLeuGlnThrGluProGlnGlyAsnProAspGlyGluGlyAspArg 50
         ||||||||||||||||||||    ||||||    |||:::|||
     320 CCTCCCTACAGACAGAACCGCAACGTAATCCGACGGGGAAGGGG...ACC 366

      51 .CysProHisGlySerProGlnGlyProLeuAlaProProAlaSerProG 67
          ||||||||||||||||||||||||||||||||||||||||||||||||||
     367 GTGCCCCCACGGCAGCCCTCAGGGCCCGCTGGCCCCACCGGCCAGCCCTG 416

      67 lyProPheAlaThrArgSerProLeuPheIlePheValArgArgSerSer 83
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     417 GCCCTTTTGCTACCAGATCCCCACTTTTCATCTTTGTGAGAAGATCTTCT 466

      84 LeuLeuSerArgSerSerSerGlyTyrPheSerPheAspThrAspArgSe 100
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     467 CTGCTGTCCCGGTCCTCCAGTGGGTATTTCTCTTTTGACACAGACAGGAG 516

     100 rProAlaProMetSerCysAspLysSerThrGlnThrProSerProProC 117
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     517 CCCGGCACCCATGAGTTGTGACAAGTCAACACAAACCCCAAGTCCTCCTT 566

     117 ysGlnAlaPheAsnHisTyrLeuSerAlaMetAlaSerIleArgGlnSer 133
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     567 GCCAGGCCTTCAACCACTATCTCAGTGCAATGGCTTCCATACGACAGTCT 616

     134 GlnGluGluProGluAspLeuArgProGluIleArgIleAlaGlnGluLe 150
         ||||||||||||||||||||||||||||||||||||||||||||||||||

     617 CAGGAGGAACCTGAAGATCTGCGCCCGGAGATACGGATTGCACAGGAGCT 666

     150 uArgArgIleGlyAspGluPheAsnGluThrTyrThrArgArgValPheA 167
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     667 GCGGCGGATCGGAGACGAGTTCAACGAAACTTACACAAGGAGGGTGTTTG 716

     167 laAsnAspTyrArgGluAlaGluAspHisProGlnMetValIleLeuGln 183
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     717 CAAATGATTACCGCGAGGCTGAAGACCACCCTCAAATGGTTATCTTACAA 766

     184 LeuLeuArgPheIlePheArgLeuValTrpArgArgHis 196
         ||||||||||||||||||||||||||||||||||||
     767 CTGTTACGCTTTATCTTCCGTCTGGTATGGAGAAGGCAT 805

seq_name: gb_est2:BG921698

seq_documentation_block:
LOCUS       BG921698      935 bp    mRNA            EST       05-JUN-2001
DEFINITION  602825518F1 NCI_CGAP_Mam6 Mus musculus cDNA clone IMAGE:4954300 5',
            mRNA sequence.
ACCESSION   BG921698
VERSION     BG921698.1  GI:14302174
KEYWORDS    EST.
SOURCE      house mouse.
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 935)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Jeffrey Green M.D.
             cDNA Library Preparation: Life Technologies, Inc.
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Incyte Genomics, Inc.
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLAM10915  row: c  column: 05
            High quality sequence start: 3
            High quality sequence stop: 786.
FEATURES             Location/Qualifiers
     source          1. .935
                     /organism="Mus musculus"
                     /strain="FVB/N"
                     /db_xref="taxon:10090"
                     /clone="IMAGE:4954300"
                     /clone_lib="NCI_CGAP_Mam6"
                     /sex="female, virgin"
                     /tissue_type="infiltrating ductal carcinoma"
                     /dev_stage="5 months"
                     /lab_host="DH10B"
                     /note="Organ: mammary; Vector: pCMV-SPORT6; Site_1: SalI;
                     Site_2: NotI; Cloned unidirectionally.  Primer: Oligo dT.
                     Library constructed by Life Technologies. Investigator
                     providing samples: Jeffrey Green, M.D., NIH"
BASE COUNT      203 a    283 c    276 g    173 t
ORIGIN


alignment_scores:
           Quality:  918.00            Length:    198
             Ratio:  4.756               Gaps:      3
 Percent Similarity:  97.475   Percent Identity: 96.465

alignment_block:
 US-09-508-832-6 x BG921698   ..

 Align_seg 1/1  to: BG921698  from: 1  to: 935

       1 MetAlaLysGlnProSerAspValSerSerGluCysAspArgGluGlyGl 17
```

```
     |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
   208 ATGGCCAAGCAACCTTCTGATGTAAGTTCTGAGTGTGACAGAGAAGGTGG 257

    17 yGlnLeuGlnProAlaGluArgProProGlnLeuArgProGlyAlaProT 34
       ||||||||||||||||||||||||||||||||||||||||||||||||||||
   258 ACAATTGCAGCCTGCTGAGAGGCCTCCCCAGCTCAGGCCTGGGGCCCCTA 307

    34 hrSerLeuGlnThrGluProGlnGlyAsnProAspGlyGluGlyAspArg 50
       ||||||||||||||||||||||||||||||||||||||||||||||||||||
   308 CCTCCCTACAGACAGAACCGCAAGGTAATCCCGACGGCGAAGGGGACCGC 357

    51 CysProHisGly.SerProGlnGlyProLeuAlaProProAlaSerProG 67
       ||||||||||||    ||||||||||||||||||||||||||||||||||
   358 TGCCCCCACGGCATGCCCTCAGGGCCCGCTGGCCCCACCGGCCAGCCCTG 407

    67 lyPro.PheAlaThrArgSerProLeuPheIlePheValArgArgSerSe 83
       |||||  ||||||||||||||||||||||||||||||||||||||||||||
   408 GCCCTTTTTGCTACCAGATCCCCACTTTTCATCTTTGTGAGAAGATCTTC 457

    83 rLeuLeuSerArgSerSerSerGlyTyrPheSerPheAspThrAspArgS 100
       ||||||||||||||||||||||||||||||||||||||||||||||||||||
   458 TCTGCTGTCCCGGTCCTCCAGTGGGTATTTCTCTTTTGACACAGACAGGA 507

   100 erProAlaProMetSerCysAspLysSerThrGlnThrProSerProPro 116
       ||||||||||||||||||||||||||||||||||    ||||||||||||
   508 GCCCGGCACCCATGAGTTGTGACAAGTCAACACAAAC.CCAAGTCCTCCT 556

   117 CysGlnAlaPheAsnHisTyrLeuSerAlaMetAlaSerIleArgGlnSe 133
       ||||||||||||||||||||||||||||||||||||||||||||||||||||
   557 TGCCAGGCCTTCAACCACTATCTCAGTGCAATGGCTTCCATACGACAGTC 606

   133 rGlnGluGluProGluAspLeuArgProGluIleArgIleAlaGlnGluL 150
       ||||||||||||||||||||||||||||||||||||||||||||||||||||
   607 TCAGGAGGAACCTGAAGATCTGCGCCCGGAGATACGGATTGCACAGGAGC 656

   150 euArgArgIleGlyAspGluPheAsnGluThrTyrThrArgArgValPhe 166
       ||||||||||||||||||||||||||||||||||||||||||||||||||||
   657 TGCGGCGGATCGGAGACGAGTTCAACGAAACTTACACAAGGAGGGTGTTT 706

   167 AlaAsnAspTyrArgGluAlaGluAspHisProGlnMetValIleLeuGl 183
       ||||||||||||||||||||||||    ||||||||||||||||||||||||
   707 GCAAATGATTACCGCGAGGCTGAAGA.CACCCTCAAATGGTTATCTTACA 755

   183 n..LeuLeuArgPheIlePheArgLeuValTrpArgArg 195
       |  :::||||||||||||||||||||||||||||   :::
   756 AACGTGTTACGCTTTATCTTCCGTCTGGTATGGGCGAAA 794
```

seq_name: gb_est2:BF021882

seq_documentation_block:

```
LOCUS       BF021882      452 bp    mRNA            EST       29-DEC-2000
DEFINITION  uy59b09.y1 McCarrey Eddy round spermatid Mus musculus cDNA clone
            IMAGE:3663833 5' similar to TR:O54918 O54918 BCL2 INTERACTING
            MEDIATOR OF CELL DEATH ;, mRNA sequence.
ACCESSION   BF021882
VERSION     BF021882.1  GI:10753214
KEYWORDS    EST.
SOURCE      house mouse.
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 452)
  AUTHORS   Marra,M., Hillier,L., Kucaba,T., Martin,J., Beck,C., Wylie,T.,
            Underwood,K., Steptoe,M., Theising,B., Allen,M., Bowers,Y., Person
            ,B., Swaller,T., Gibbons,M., Pape,D., Harvey,N., Schurk,R., Ritter
            ,E., Kohn,S., Shin,T., Jackson,Y., Cardenas,M., McCann,R.,
            Waterston,R. and Wilson,R.
  TITLE     The WashU-NCI Mouse EST Project 1999
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Marra M/WashU-NCI Mouse EST Project 1999
            Washington University School of Medicine
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: mouseest@watson.wustl.edu
            This clone is available royalty-free through LLNL ; contact the
            IMAGE Consortium (info@image.llnl.gov) for further information.
            MGI:1424601
            Seq primer: Primer name ambiguous
            High quality sequence stop: 386.
FEATURES             Location/Qualifiers
     source          1. .452
                     /organism="Mus musculus"
                     /strain="CD-1"
                     /db_xref="taxon:10090"
                     /clone="IMAGE:3663833"
                     /clone_lib="McCarrey Eddy round spermatid"
                     /sex="male"
                     /tissue_type="round spermatids, pooled from multiple mice"
                     /dev_stage="60 day"
                     /lab_host="DH10B (phage-resistant)"
                     /note="Organ: testis; Vector: pBluescript SK+ (Stratagene
                     ); Site_1: XhoII; Site_2: EcoRI; cDNA oligo dT-primed
                     [5'-(GA)10-ACTAGTCTCGAGTTTTTTTTTTTTT-3'] and directionally
                     cloned  using 5' linkers 5'-AATTCGGCACGAG-3' and
                     5'-CTCGTGCCG-3'. Size selection of >400bp material gives
                     average insert size ranging from 1-2 kb. Library was mass
                     excised (from  lambda-UniZAP-XR) and resulting
                     single-stranded phagemids were prepped and tranformed
                     into DH10B. Library contains 98.5% recombinants.
                     References: J. Androl. 20:635-639 and Gene 25:263-269.
                     Library constructed and donated by J. McCarrey, Ph.D.
                     (Southwest Foundation for Biomedical Research, Dept. of
                     Genetics); excision done by E.M. Eddy, Ph.D. (National
                     Institutes of  Health, National Institute of Environmental
                     Health Sciences).  Original lambda-based library is
                     available through ATCC, catalog #63423."
BASE COUNT      106 a    130 c    112 g    104 t
ORIGIN
```

alignment_scores:

Quality: 730.00 Length: 139
Ratio: 5.252 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
US-09-508-832-6 x BF021882 ..

Align seg 1/1 to: BF021882 from: 1 to: 452

```
    58 GlyProLeuAlaProProAlaSerProGlyProPheAlaThrArgSerPr 74
       ||||||||||||||||||||||||||||||||||||||||||||||||||||
     1 GGCCCGCTGGCCCCACCGGCCAGCCCTGGCCCTTTTGCTACCAGATCCCC 50

    74 oLeuPheIlePheValArgArgSerSerLeuLeuSerArgSerSerSerG 91
       ||||||||||||||||||||||||||||||||||||||||||||||||||||
    51 ACTTTTCATCTTTGTGAGAAGATCTTCTCTGCTGTCCCGGTCCTCCAGTG 100

    91 lyTyrPheSerPheAspThrAspArgSerProAlaProMetSerCysAsp 107
       ||||||||||||||||||||||||||||||||||||||||||||||||||||
   101 GGTATTTCTCTTTTGACACAGACAGGAGCCCGGCACCCATGAGTTGTGAC 150

   108 LysSerThrGlnThrProSerProProCysGlnAlaPheAsnHisTyrLe 124
       ||||||||||||||||||||||||||||||||||||||||||||||||||||
   151 AAGTCAACACAAACCCCAAGTCCTCCTTGCCAGGCCTTCAACCACTATCT 200

   124 uSerAlaMetAlaSerIleArgGlnSerGlnGluGluProGluAspLeuA 141
       ||||||||||||||||||||||||||||||||||||||||||||||||||||
   201 CAGTGCAATGGCTTCCATACGACAGTCTCAGGAGGAACCTGAAGATCTGC 250

   141 rgProGluIleArgIleAlaGlnGluLeuArgArgIleGlyAspGluPhe 157
       ||||||||||||||||||||||||||||||||||||||||||||||||||||
   251 GCCCGGAGATACGGATTGCACAGGAGCTGCGGCGGATCGGAGACGAGTTC 300

   158 AsnGluThrTyrThrArgArgValPheAlaAsnAspTyrArgGluAlaGl 174
```

```
     ||||||||||||||||||||||||||||||||||||||||||||||||
 301 AACGAAACTTACACAAGGAGGGTGTTTGCAAATGATTACCGCGAGGCTGA 350

 174 uAspHisProGlnMetValIleLeuGlnLeuLeuArgPheIlePheArgL 191
     ||||||||||||||||||||||||||||||||||||||||||||||||||
 351 AGACCACCCTCAAATGGTTATCTTACAACTGTTACGCTTTATCTTCCGTC 400

 191 euValTrpArgArgHis 196
     ||||||||||||||||
 401 TGGTATGGAGAAGGCAT 417

seq_name: gb_gss:AZ706148

seq_documentation_block:
LOCUS       AZ706148      580 bp    DNA             GSS       24-JAN-2001
DEFINITION  RPCI-23-227P3.TV RPCI-23 Mus musculus genomic clone RPCI-23-227P3,
            DNA sequence.
ACCESSION   AZ706148
VERSION     AZ706148.1  GI:12433319
KEYWORDS    GSS.
SOURCE      house mouse.
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 580)
  AUTHORS   Zhao,S., Nierman,W., Feldblyum,T., Malek,J., Shatsman,S., Akinret
            ,B., Levins,M., Mcgann,S., Tsegaye,G., Geer,K., Krol,M., de Jong,P.
            and Fraser,C.M.
  TITLE     Mouse BAC End Sequences from Library RPCI-23
  JOURNAL   Unpublished (1999)
COMMENT     Other_GSSs: RPCI-23-227P3.TJ
            Contact: Shaying Zhao
            Department of Eukaryotic Genomics
            The Institute for Genomic Research
            9712 Medical Center Dr., Rockville, MD 20850, USA
            Tel: 301 838 0200
            Fax: 301 838 0208
            Email: szhao@tigr.org
            Clones are derived from the mouse BAC library RPCI-23. For BAC
            library availability, please contact Pieter de Jong
            (pdejong@mail.cho.org). Clones may be purchased from BACPAC
            Resources (http://www.chori.org/bacpac/orderingframe.htm). BAC end
            page: http://ww .tigr.org/tdb/bac_ends/mouse/bac_end_intro.html
            Plate: 227  row: P  column: 3
            Seq primer: T7
            Class: BAC ends.
FEATURES             Location/Qualifiers
     source          1. .580
                     /organism="Mus musculus"
                     /strain="C57BL/6J"
                     /db_xref="taxon:10090"
                     /clone="RPCI-23-227P3"
                     /clone_lib="RPCI-23"
                     /sex="Female"
                     /lab_host="DH10B"
                     /note="Organ: Kidney/Brain; Vector: pBACe3.6; Site_1:
                     EcoRI; Site_2: EcoRI; Female C57BL/6J mouse kidney and/or
                     brain genomic DNA was isolated and  partially digested
                     with a combination of EcoRI and EcoRI Methylase.   Size
                     selected DNA was cloned into the pBACe3.6 vector at the
                     EcoRI  sites.  The ligation products were transformed into
                     DH10B  electrocompetent cells (BRL Life Technologies).   "
BASE COUNT      138 a    162 c    138 g    142 t
ORIGIN


alignment_scores:
            Quality:  689.00            Length:    127
              Ratio:  5.425               Gaps:      0
 Percent Similarity: 100.000   Percent Identity: 100.000

alignment_block:
 US-09-508-832-6 x AZ706148   ..

 Align seg 1/1  to: AZ706148  from: 1  to: 580

       1 MetAlaLysGlnProSerAspValSerSerGluCysAspArgGluGlyGl 17
         ||||||||||||||||||||||||||||||||||||||||||||||||||
      90 ATGGCCAAGCAACCTTCTGATGTAAGTTCTGAGTGTGACAGAGAAGGTGG 139

      17 yGlnLeuGlnProAlaGluArgProProGlnLeuArgProGlyAlaProT 34
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     140 ACAATTGCAGCCTGCTGAGAGGCCTCCCCAGCTCAGGCCTGGGGCCCCTA 189

      34 hrSerLeuGlnThrGluProGlnGlyAsnProAspGlyGluGlyAspArg 50
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     190 CCTCCCTACAGACAGAACCGCAAGGTAATCCCGACGGCGAAGGGGACCGC 239

      51 CysProHisGlySerProGlnGlyProLeuAlaProProAlaSerProGl 67
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     240 TGCCCCCACGGCAGCCCTCAGGGCCCGCTGGCCCCACCGGCCAGCCCTGG 289

      67 yProPheAlaThrArgSerProLeuPheIlePheValArgArgSerSerL 84
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     290 CCCTTTTGCTACCAGATCCCCACTTTTCATCTTTGTGAGAAGATCTTCTC 339

      84 euLeuSerArgSerSerSerGlyTyrPheSerPheAspThrAspArgSer 100
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     340 TGCTGTCCCGGTCCTCCAGTGGGTATTTCTCTTTTGACACAGACAGGAGC 389

     101 ProAlaProMetSerCysAspLysSerThrGlnThrProSerProProCy 117
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     390 CCGGCACCCATGAGTTGTGACAAGTCAACACAAACCCCAAGTCCTCCTTG 439

     117 sGlnAlaPheAsnHisTyrLeuSerAlaMet 127
         |||||||||||||||||||||||||||||||
     440 CCAGGCCTTCAACCACTATCTCAGTGCAATG 470

seq_name: gb_est2:BG173095

seq_documentation_block:
LOCUS       BG173095      668 bp    mRNA            EST       06-FEB-2001
DEFINITION  602336666F1 NCI_CGAP_Mam1 Mus musculus cDNA clone IMAGE:4459720 5',
            mRNA sequence.
ACCESSION   BG173095
VERSION     BG173095.1  GI:12679707
KEYWORDS    EST.
SOURCE      house mouse.
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 668)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Gilbert Smith, Ph.D.
            cDNA Library Preparation: Life Technologies, Inc.
            cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
            DNA Sequencing by: Incyte Genomics, Inc.
            Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLAM10260  row: c  column: 17
            High quality sequence stop: 599.
FEATURES             Location/Qualifiers
     source          1. .668
                     /organism="Mus musculus"
                     /strain="FVB/N"
                     /db_xref="taxon:10090"
                     /clone="IMAGE:4459720"
                     /clone_lib="NCI_CGAP_Mam1"
                     /tissue_type="tumor, biopsy sample"
                     /dev_stage="3 months, virgin"
                     /lab_host="DH10B"
```

```
                    /note="Organ: mammary; Vector: pCMV-SPORT6; Site_1: SalI;
                    Site_2: NotI; Cloned unidirectionally.  Primer: Oligo dT.
                    Library constructed by Life Technologies. Investigator
                    providing samples: Gilbert Smith, NIH"
BASE COUNT      135 a    235 c    162 g    136 t
ORIGIN


alignment_scores:
            Quality:  606.50             Length:    150
              Ratio:  4.460                Gaps:      6
 Percent Similarity:  90.667   Percent Identity: 85.333

alignment_block:
 US-09-508-832-6 x BG173095   ..

 Align seg 1/1  to: BG173095  from: 1  to: 668

       1 MetAlaLysGlnProSerAspValSerSerGluCysAspArgGluGlyGl 17
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     223 ATGGCCAAGCAACCTTCTGATGTAAGTTCTGAGTGTGACAGAGAAGGTGG 272

      17 yGlnLeuGlnProAlaGluArgProProGlnLeuArgProGlyAlaProT 34
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     273 ACAATTGCAGCCTGCTGAGAGGCCTCCCCAGCTCAGGCCTGGGGCCCCTA 322

      34 hrSerLeuGlnThrGluProGlnGlyAsnProAspGlyGluGlyAspArg 50
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     323 CCTCCCTACAGACAGAACCGCAAGGTAATCCCGACGGCGAAGGGGACCTG 372

      51 .CysProHisGlySerProGlnGlyProLeuAlaProProAlaSerProG 67
         |||||||||||||||||||||||||||||:::|||   :::|||||||
     373 CTGCCCCCACGGCAGCCCTCAGGGCCCGCTTGGCCCCACCGGCAGCCCTG 422

      67 lyProPheAlaThrArgSer.ProLeuPheIlePheValArgArgSerSe 83
         |||||||||||||||||||| ||||||||||||||||||||||||||||||
     423 GCCCTTTTGCTACCAGATCCCCCACTTTTCATCTTTGTGAGAAGATCTTC 472

      83 rLeuLeuSerArgSerSerSerGlyTyrPheSer...PheAspThrAspA 99
         ||||||||||||||||||:::      :::|||   |||||||||   :
     473 TCTGCTGTCCCGGTCCTCCCAGTGGGTATATTCTCCTTTTGACACAGCAC 522

      99 rgSerProAlaProMetSerCysAspLysSerThrGln.ThrProSerPr 115
         :::::|||:::      :::||||||||||||||||| ||||||||||
     523 AGGAGCCCGGCACCCATGAGTGTGACAAGTCAACACAAAACCCCAAGTCC 572

     115 oProCysGlnAlaPheAsnHisTyrLeuSer.AlaMetAlaSerIleArg 131
         |||||||||||||||||||||||||||||| |||||||||   ||||||
     573 TCCTTGCCAGGCCTTCAACCACTATCTCAGTTGCAATGGCTTTCATACGA 622

     132 GlnSer.GlnGluGluProGluAspLeuArgProGluIle 144
         |||||| |||||||||||||||||||||||||||||||
     623 CAGTCTCCAGGAGGAACCTGAGGATCTGCGCCCGGAGATC 662

seq_name: gb_est1:AI971169

seq_documentation_block:
LOCUS       AI971169      492 bp    mRNA            EST       08-MAR-2000
DEFINITION  wr24h12.x1 NCI_CGAP_Pr28 Homo sapiens cDNA clone IMAGE:2488679 3'
            similar to TR:O43522 O43522 BIML. [1] ;, mRNA sequence.
ACCESSION   AI971169
VERSION     AI971169.1  GI:5767995
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 492)
  AUTHORS   NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
  TITLE     National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
            Tumor Gene Index
  JOURNAL   Unpublished (1997)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Michael J. Brownstein, M.D., Ph.D., Michael R.
            Emmert-Buck, M.D., Ph.D.
            cDNA Library Preparation: M. Bento Soares, Ph.D.
            cDNA Library Arrayed by: Greg Lennon, Ph.D.
            DNA Sequencing by: Washington University Genome Sequencing Center
            Clone distribution: NCI-CGAP clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            www-bio.llnl.gov/bbrp/image/image.html
            Insert Length: 712   Std Error: 0.00
            Seq primer: -40UP from Gibco
            High quality sequence stop: 450.
FEATURES             Location/Qualifiers
     source          1. .492
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:2488679"
                     /clone_lib="NCI_CGAP_Pr28"
                     /sex="male"
                     /dev_stage="adult"
                     /lab_host="DH10B"
                     /note="Organ: prostate; Vector: pT7T3D-Pac (Pharmacia)
                     with a modified polylinker; Plasmid DNA from the
                     normalized library NCI_CGAP_Pr22 was prepared, and ss
                     circles were made in vitro. Following HAP purification,
                     this DNA was used as tracer in a subtractive hybridization
                     reaction. The driver was PCR-amplified cDNAs from a pool
                     of 5,000 clones made from the same library (cloneIDs
                     985608-986759, 1101192-1101959, and 1217928-1220615).
                     Subtraction by Bento Soares and M. Fatima Bonaldo.  "
BASE COUNT      119 a    107 c    130 g    134 t      2 others
ORIGIN


alignment_scores:
            Quality:  332.00             Length:    130
              Ratio:  4.676                Gaps:      1
 Percent Similarity:  54.615   Percent Identity: 53.846

alignment_block:
 US-09-508-832-6 x AI971169/rev   ..

 Align seg 1/1  to reverse of: AI971169  from: 1  to: 492

       1 MetAlaLysGlnProSerAspValSerSerGluCysAspArgGluGlyGl 17
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     425 ATGGCAAAGCAACCTTCTGATGTAAGTTCTGAGTGTGACCGAGAAGGTAG 376

      17 yGlnLeuGlnProAlaGluArgProProGlnLeuArgProGlyAlaProT 34
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     375 ACAATTGCAGCCTGCGGAGAGGCCTCCCCAGCTCAGACCTGGGGCCCCTA 326

      34 hrSerLeuGlnThrGluProGlnGlyAsnProAspGlyGluGlyAspArg 50
         ||||||||||||||||||||||||
     325 CCTCCCTACAGACAGAGCCACAA..... 303

      51 CysProHisGlySerProGlnGlyProLeuAlaProProAlaSerProGl 67

     303 ..... 303

      67 yProPheAlaThrArgSerProLeuPheIlePheValArgArgSerSerL 84

     303 ..... 303

      84 euLeuSerArgSerSerSerGlyTyrPheSerPheAspThrAspArgSer 100
                                                 |||||||||
     302 .....GACAGGAGC 294

     101 ProAlaProMetSerCysAspLysSerThrGlnThrProSerProProCy 117
         ||||||||||||||||||||||||||||||||||   ||||||||||||
     293 CCAGCACCCATGAGTTGTGACAAATCAACACAAACNNCAAGTCCTCCTTG 244

     117 sGlnAlaPheAsnHisTyrLeuSerAlaMetAlaSerIle 130
```

```
          ||||||||||||||||||||||||||||||:::   |||
     243 CCAGGCCTTCAACCACTATCTCAGTGCAATGGTAGTCATC 204

seq_name: gb_est2:BF319454

seq_documentation_block:
LOCUS       BF319454      389 bp    mRNA            EST       29-DEC-2000
DEFINITION  uy59b09.x1 McCarrey Eddy round spermatid Mus musculus cDNA clone
            IMAGE:3663833 3' similar to TR:O54918 O54918 BCL2 INTERACTING
            MEDIATOR OF CELL DEATH ;, mRNA sequence.
ACCESSION   BF319454
VERSION     BF319454.1  GI:11268195
KEYWORDS    EST.
SOURCE      house mouse.
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 389)
  AUTHORS   Marra,M., Hillier,L., Kucaba,T., Martin,J., Beck,C., Wylie,T.,
            Underwood,K., Steptoe,M., Theising,B., Allen,M., Bowers,Y., Person
            ,B., Swaller,T., Gibbons,M., Pape,D., Harvey,N., Schurk,R., Ritter
            ,E., Kohn,S., Shin,T., Jackson,Y., Cardenas,M., McCann,R.,
            Waterston,R. and Wilson,R.
  TITLE     The WashU-NCI Mouse EST Project 1999
  JOURNAL   Unpublished (1999)
COMMENT     Other_ESTs: uy59b09.y1
            Contact: Marra M/WashU-NCI Mouse EST Project 1999
            Washington University School of Medicine
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: mouseest@watson.wustl.edu
            This clone is available royalty-free through LLNL ; contact the
            IMAGE Consortium (info@image.llnl.gov) for further information.
            MGI:1424601
            High quality sequence stop: 325.
FEATURES             Location/Qualifiers
     source          1. .389
                     /organism="Mus musculus"
                     /strain="CD-1"
                     /db_xref="taxon:10090"
                     /clone="IMAGE:3663833"
                     /clone_lib="McCarrey Eddy round spermatid"
                     /sex="male"
                     /tissue_type="round spermatids, pooled from multiple mice"
                     /dev_stage="60 day"
                     /lab_host="DH10B (phage-resistant)"
                     /note="Organ: testis; Vector: pBluescript SK+ (Stratagene
                     ); Site_1: XhoII; Site_2: EcoRI; cDNA oligo dT-primed
                     [5'-(GA)10-ACTAGTCTCGAGTTTTTTTTTTTT-3'] and directionally
                     cloned  using 5' linkers 5'-AATTCGGCACGAG-3' and
                     5'-CTCGTGCCG-3'. Size selection of >400bp material gives
                     average insert size ranging from 1-2 kb. Library was mass
                     excised (from  lambda-UniZAP-XR) and resulting
                     single-stranded phagemids were prepped and tranformed
                     into DH10B. Library contains 98.5% recombinants.
                     References: J. Androl. 20:635-639 and Gene 25:263-269.
                     Library constructed and donated by J. McCarrey, Ph.D.
                     (Southwest Foundation for Biomedical Research, Dept. of
                     Genetics); excision done by E.M. Eddy, Ph.D. (National
                     Institutes of  Health, National Institute of Environmental
                     Health Sciences).  Original lambda-based library is
                     available through ATCC, catalog #63423."
BASE COUNT      100 a    104 c     89 g     96 t
ORIGIN


alignment_scores:
           Quality:  249.00            Length:     53
             Ratio:  4.882               Gaps:      1
 Percent Similarity:  96.226   Percent Identity: 94.340

alignment_block:
```

```
US-09-508-832-6 x BF319454/rev   ..

Align seg 1/1  to reverse of: BF319454  from: 1  to: 389

    145 ArgIleAlaGlnGluLeuArgArgIleGly.AspGluPheAsnGluThrT 161
        |||:::|||||||||||||||   |||||| ||||||||||||||||||
    389 CGATTTGCACAGGAACTTCGGGGGATCGGAAGACGAGTTCAACGAAACTT 340

    161 yrThrArgArgValPheAlaAsnAspTyrArgGluAlaGluAspHisPro 177
        ||||||||||||||||||||||||||||||||||||||||||||||||||
    339 ACACAAGGAGGGTGTTTGCAAATGATTACCGCGAGGCTGAAGACCACCCT 290

    178 GlnMetValIleLeuGlnLeuLeuArgPheIlePheArgLeuValTrpAr 194
        ||||||||||||||||||||||||||||||||||||||||||||||||||
    289 CAAATGGTTATCTTACAACTGTTACGCTTTATCTTCCGTCTGGTATGGAG 240

    194 gArgHis 196
        |||||||
    239 AAGGCAT 233

seq_name: gb_est1:AW629314

seq_documentation_block:
LOCUS       AW629314      664 bp    mRNA.           EST       31-MAR-2000
DEFINITION  hi56e02.x1 Soares_NFL_T_GBC_S1 Homo sapiens cDNA clone
            IMAGE:2976314 3' similar to TR:O43522 O43522 BIML. [1] ;, mRNA
            sequence.
ACCESSION   AW629314
VERSION     AW629314.1  GI:7376104
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 664)
  AUTHORS   NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
  TITLE     National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
            Tumor Gene Index
  JOURNAL   Unpublished (1997)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            This clone is available royalty-free through LLNL ; contact the
            IMAGE Consortium (info@image.llnl.gov) for further information.
            Seq primer: -40UP from Gibco
            High quality sequence stop: 458.
FEATURES             Location/Qualifiers
     source          1. .664
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:2976314"
                     /clone_lib="Soares_NFL_T_GBC_S1"
                     /lab_host="DH10B"
                     /note="Organ: pooled; Vector: pT7T3D-Pac (Pharmacia) with
                     a modified polylinker; Site_1: Not I; Site_2: Eco RI;
                     Equal amounts of plasmid DNA from three normalized
                     libraries (fetal lung NbHL19W, testis NHT, and B-cell
                     NCI_CGAP_GCB1) were mixed, and ss circles were made in
                     vitro. Following HAP purification, this DNA was used as
                     tracer in a subtractive hybridization reaction. The driver
                     was PCR-amplified cDNAs from pools of 5,000 clones made
                     from the same 3 libraries. The pools consisted of
                     I.M.A.G.E. clones 297480-302087, 682632-687239,
                     726408-728711, and 729096-731399. Subtraction by Bento
                     Soares and M. Fatima Bonaldo. "
BASE COUNT      176 a    131 c    148 g    208 t      1 others
ORIGIN


alignment_scores:
           Quality:  249.00            Length:    125
             Ratio:  4.220               Gaps:      1
 Percent Similarity:  47.200   Percent Identity: 43.200
```

```
alignment_block:
 US-09-508-832-6 x AW629314/rev   ..

 Align seg 1/1  to reverse of: AW629314  from: 1  to: 664

       3 LysGlnProSerAspValSerSerGluCysAspArgGluGlyGlyGlnLe 19
         |||   |||   |||   |||   ||||||      ||||||   |||::
     664 AAGCCACCTTTTGATGAAAGTTTTGAGTGTGCCNGAGAAGGTAGACAAAT 615

      19 uGlnProAlaGluArgProProGlnLeuArgProGlyAlaProThrSerL 36
         :|||||||||:::|||   ||||||:::|||   ||||||||||||||||
     614 GCAGCCTGCGAAGAGGCTTCCCCAGTTCAGACTTGGGGCCCCTACCTCCT 565

      36 euGlnThrGluProGlnGlyAsnProAspGlyGluGlyAspArgCysPro 52
           :::||||||||||||
     564 ACAAGACAGAGCCACAA..................................... 548

      53 HisGlySerProGlnGlyProLeuAlaProProAlaSerProGlyProPh 69

     548 ..................................................... 548

      69 eAlaThrArgSerProLeuPheIlePheValArgArgSerSerLeuLeuS 86

     548 ..................................................... 548

      86 erArgSerSerSerGlyTyrPheSerPheAspThrAspArgSerProAla 102
                                          ||||||||||||||||
     547 ..................................GACAGGAGCCCAGCA 533

     103 ProMetSerCysAspLysSerThrGlnThrProSerProProCysGlnAl 119
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     532 CCCATGAGTTGTGACAAATCAACACAAACCCCAAGTCCTCCTTGCCAGGC 483

     119 aPheAsnHisTyrLeuSerAlaMet 127
         |||||||||||||||||||:::|||
     482 CTTCAACCACTATCTCAGTGTAATG 458

seq_name: gb_est1:AF209718

seq_documentation_block:
LOCUS       AF209718      537 bp    mRNA            EST       30-MAY-2000
DEFINITION  AF209718 Xenopus laevis intestine adult Xenopus laevis cDNA clone
            pXLG10 similar to Mus musculus BimEL, mRNA sequence.
ACCESSION   AF209718
VERSION     AF209718.1  GI:8110110
KEYWORDS    EST.
SOURCE      African clawed frog.
  ORGANISM  Xenopus laevis
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidea; Pipidae;
            Xenopodinae; Xenopus.
REFERENCE   1  (bases 1 to 537)
  AUTHORS   Minter,R.
  TITLE     Development of Antibody Technology to Identify Natural Killer Cell
            Surface Antigens in Xenopus laevis. Thesis (1999) University of
            Durham, South Rd., Durham, UK
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Minter,R., Horton,J.D. and Watson,M.D.
            Biological Sciences
            University of Durham
            South Road, Durham, DH1 3LE, UK
            Email: martin.watson@durham.ac.uk.
FEATURES             Location/Qualifiers
     source          1. .537
                     /organism="Xenopus laevis"
                     /db_xref="taxon:8355"
                     /clone="pXLG10"
                     /clone_lib="Xenopus laevis intestine adult"
                     /tissue_type="intestine"
                     /cell_type="epithelial lymphocyte"
                     /dev_stage="adult"
BASE COUNT      126 a    129 c    143 g    139 t
ORIGIN

alignment_scores:
           Quality:  202.00            Length:    107
             Ratio:  2.730               Gaps:      3
 Percent Similarity:  69.159   Percent Identity: 50.467

alignment_block:
 US-09-508-832-6 x AF209718   ..

 Align seg 1/1  to: AF209718  from: 1  to: 537

       1 MetAlaLysGlnProSerAspValSerSerGluCysAspArg...GluGl 16
         ||||||||||||||||||   :::|||   ||||||:::      |||||
     223 ATGGCCAAACAACCGTCGGTCTTGAGTCCGGAGTGTAATAGTGGTGAAGG 272

      16 yGlyGlnLeuGlnProAlaGluArgProProGlnLeuArgPro....... 30
         ||||||||||||||   ::::::|||      :::   ||||||
     273 TGGCCAGTTACAATCAACAAGCAGACGACATTCTCATCGTCCTCGCAGAA 322

      31 ..GlyAlaProThrSerLeuGlnThrGluProGlnGlyAsnProAspGly 46
           ||||||||||||||||||::::::      |||||||||   :::
     323 GAGGGGCCCCCACCTCTCTTAGCAGTCCTTTTCAAGGTAATCAATCAGAT 372

      47 GluGlyAspArgCysProHisGlySerProGlnGlyProLeuAlaProPr 63
         ||||||  .          :::::|||   ||||||         ||
     373 GAGGGTGGGAGCTCCTCAGCCAGCACTCCTTGGGGTCCTACTTTATCGCC 422

      63 oAlaSerProGlyProPheAlaThrArgSerProLeuPheIlePheValA 80
         |   ||||||:::   |||::::::|||||||||      ::::::||||
     423 TTATAGCCCCAGTTCCTTTGTCAACAGATCACCCCATTGCATGCTGGTAA 472

      80 rgArgSerSerLeuLeuSerArgSerSerSerGlyTyrPheSerPheAsp 96
         ||   |||||||||:::|||:::   ||||||||||||||:::|||
     473 GAGGATCATCACTTGTCTCAAAACC.TCAAGTGGCTATTTTACATTC... 518

      97 ThrAspArgSerProAlaPro 103
         :::  ||||||:::|||
     519 ...GAAGGGAGTCCTGGGCCT 536

seq_name: gb_est1:AW748960

seq_documentation_block:
LOCUS       AW748960      157 bp    mRNA            EST       28-APR-2000
DEFINITION  RC4-BT0312-081199-011-b02 BT0312 Homo sapiens cDNA, mRNA sequence.
ACCESSION   AW748960
VERSION     AW748960.1  GI:7663892
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 157)
  AUTHORS   Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R.,
            Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F.,
            Goldman,G.H., Carvalho,A.F., Matsukuma,A., Baia,G.S., Simpson,D.H.,
            Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V., O'Hare
            ,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and
            Simpson,A.J.
  TITLE     Shotgun sequencing of the human transcriptome with ORF expressed
            sequence tags
  JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
  MEDLINE   20202663
COMMENT     Contact: Simpson A.J.G.
            Laboratory of Cancer Genetics
            Ludwig Institute for Cancer Research
            Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
            Brazil
            Tel: +55-11-2704922
            Fax: +55-11-2707001
            Email: asimpson@ludwig.org.br
            This sequence was derived from the FAPESP/LICR Human Cancer Genome
            Project. This entry can be seen in the following URL
```

```
            (http://www.ludwig.org.br/scripts/gethtml2.pl?t1=RC4&t2=RC4-BT0312-
            081199-011-b02&t3=1999-11-08&t4=1)
            Seq primer: puc 18 forward
            High quality sequence start: 24
            High quality sequence stop: 45.
FEATURES              Location/Qualifiers
     source           1. .157
                      /organism="Homo sapiens"
                      /db_xref="taxon:9606"
                      /clone_lib="BT0312"
                      /dev_stage="Adult"
                      /note="Organ: breast; Vector: puc18; Site_1: SmaI; Site_2:
                      SmaI; A mini-library was made by cloning products derived
                      from  ORESTES PCR (U.S. Letters Patent application No. 196
                      ,716 -  Ludwig Institute for Cancer Research) profiles
                      into  the pUC 18 vector. Reverse transcription of tissue
                      mRNA and cDNA amplification were performed under low
                      stringency conditions."
BASE COUNT       38 a      48 c      28 g      43 t
ORIGIN


alignment_scores:
            Quality:  149.00              Length:      50
              Ratio:  3.386                 Gaps:       3
 Percent Similarity:  88.000    Percent Identity: 80.000

alignment_block:
 US-09-508-832-6 x AW748960    ..

 Align seq 1/1  to: AW748960  from: 1  to: 157

      68 ProPheAlaThrArgSerProLeuPheIlePheValArgArgSerSerLe 84
         ||||||   ::::::   ||||||||||||||||:::||||||   |||||
      15 CCTTTT...GTACAGATCCCGCTTTTCATCTTTATAAGAAGATC.TCCCT 60

      84 uLeuSerArgSerSerSerGlyTyrPheSerPheAspThrAspArgSer. 100
         ||||||||||||||||||||||||||||||||||||||||||||||||||
      61 GCTGTCACGATCCTCCAGTGGGTATTTCTCTTTTGACACAGACAGGAGCC 110

     101 ProAlaProMetSerCysAspLysSerThrGlnThrProSerProPro 116
           :::   ||||||||||||||||||   ||||||||||||||||||
     111 AGCACC...ATGAGTTGTGACAAATCACCACAAACCCCAAGTCCTCCT 155

seq_name: gb_est1:AA629308

seq_documentation_block:
LOCUS        AA629308      501 bp    mRNA            EST       16-OCT-1997
DEFINITION   zu84g06.s1 Soares_testis_NHT Homo sapiens cDNA clone IMAGE:744730
             3', mRNA sequence.
ACCESSION    AA629308
VERSION      AA629308.1  GI:2541695
KEYWORDS     EST.
SOURCE       human.
  ORGANISM   Homo sapiens
             Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
             Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE    1  (bases 1 to 501)
  AUTHORS    Hillier,L., Allen,M., Bowles,L., Dubuque,T., Geisel,G., Jost,S.,
             Krizman,D., Kucaba,T., Lacy,M., Le,N., Lennon,G., Marra,M., Martin
             ,J., Moore,B., Schellenberg,K., Steptoe,M., Tan,F., Theising,B.,
             White,Y., Wylie,T., Waterston,R. and Wilson,R.
  TITLE      WashU-NCI human EST Project
  JOURNAL    Unpublished (1997)
COMMENT      Contact: Wilson RK
             Washington University School of Medicine
             4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
             Tel: 314 286 1800
             Fax: 314 286 1810
             Email: est@watson.wustl.edu
             This clone is available royalty-free through LLNL ; contact the
             IMAGE Consortium (info@image.llnl.gov) for further information.
             Seq primer: -40m13 fwd. ET from Amersham
             High quality sequence stop: 471.
FEATURES              Location/Qualifiers
     source           1. .501
                      /organism="Homo sapiens"
                      /db_xref="GDB:5932418"
                      /db_xref="taxon:9606"
                      /clone="IMAGE:744730"
                      /clone_lib="Soares_testis_NHT"
                      /sex="male"
                      /lab_host="DH10B"
                      /note="Vector: pT7T3D-Pac (Pharmacia) with a modified
                      polylinker; Site_1: Not I; Site_2: Eco RI; 1st strand cDNA
                      was prepared from mRNA obtained from Clontech Laboratories
                      , Inc., and primed with  a Not I - oligo(dT) primer [5'
                      TGTTACCAATCTGAAGTGGGAGCGGCCGCCCAATTTTTTTTTTTTTTTT 3']. 
                      Double-stranded cDNA was ligated to Eco RI adaptors
                      (Pharmacia), digested with Not I and cloned into the Not I
                      and Eco RI sites of the modified pT7T3 vector. Library
                      went through one round of normalization to Cot5, and was
                      constructed  by Bento Soares and M. Fatima Bonaldo. "
BASE COUNT      155 a     112 c      97 g     137 t
ORIGIN


alignment_scores:
            Quality:  127.00              Length:      32
              Ratio:  4.536                 Gaps:       0
 Percent Similarity:  87.500    Percent Identity: 71.875

alignment_block:
 US-09-508-832-6 x AA629308/rev    ..

 Align seq 1/1  to reverse of: AA629308  from: 1  to: 501

     165 ValPheAlaAsnAspTyrArgGluAlaGluAspHisProGlnMetValIl 181
         ||||||   |||:::|||:::   |||||||||||||||:::||||||||
     501 GTATTTTTGAATAATTACCAAGCAGCCGAAGACCACCCACGAATGGTTAT 452

     181 eLeuGlnLeuLeuArgPheIlePheArgLeuValTrpArgArgHis 196
         ||||:::||||||||||:::|||   |||||||||||||||   |||
     451 CTTACGACTGTTACGTTACATTGTCCGCCTGGTGTGGAGAATGCAT 406

seq_name: gb_est2:BF172831

seq_documentation_block:
LOCUS        BF172831      210 bp    mRNA            EST       23-MAR-2001
DEFINITION   PCL5805 Myeloma (PCL) cDNA library Homo sapiens cDNA, mRNA
             sequence.
ACCESSION    BF172831
VERSION      BF172831.1  GI:13439045
KEYWORDS     EST.
SOURCE       human.
  ORGANISM   Homo sapiens
             Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
             Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE    1  (bases 1 to 210)
  AUTHORS    Claudio,J.O., Tang,H., Khan,E.M., Voralia,M., Li,Z., Cukerman,E.,
             Franciso-Pabalan,O., Liew,C.C. and Stewart,A.K.
  TITLE      The transcriptional phenotype of myeloma cells
  JOURNAL    Unpublished (2000)
COMMENT      Contact: A. Keith Stewart, M.D.
             Oncology Research
             University Health Network
             610 University Ave., 5-126, Toronto, Ontario, M5G 2M9, Canada
             Tel: (416) 946-4639
             Fax: (416) 946-6546
             Email: k.stewart@utoronto.ca
             PCR PRimers
             FORWARD: 5'-GCCAAGCTCGAAATTAACCCTCACTAAAGGG-3'
             BACKWARD: 5'-CCAGTGAATTGTAATACGACTCACTATAGGGCG-3'
             Seq primer: 5'-GAAATTAACCCTCACTAAAGG-3'.
FEATURES              Location/Qualifiers
     source           1. .210
```

```
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone_lib="Myeloma (PCL) cDNA library"
                     /sex="male"
                     /tissue_type="Blood"
                     /cell_type="myeloma"
                     /dev_stage="Plasma cell leukemia"
                     /note="Vector: Lambda Zap Express; Site_1: EcoRI; Site_2:
                     XhoI; mRNA was purified from plasma cell leukemia
                     patient's peripheral blood  containing >95% myeloma. An
                     oligo d(T)18 primer containing XhoI restriction  site was
                     used to prime first strand synthesis using M-MLV reverse
                     transcriptase. To protect the cDNAs from XhoI digestion in
                     subsequent  cloning step, the nucleotide analogue
                     5-methyl-dCTP was added to the  nucleotide mixture and
                     [a-32P]dATP was added to monitor the quantity and  quality
                     of first strand synthesis. After second-strand synthesis
                     and  blunting of cDNA termini, EcoRI adapters were ligated
                     , followed by kinase  treatment and digestion with XhoI.
                     The cDNAs were then size-fractionated  using Sephacryl
                     S-500 column and then ligated into EcoRI and XhoI digested
                     Lambda Zap Express vector. The ligation product was
                     packaged using Gigapack  II packaging extract. The library
                     had primary titre of approx. 1x106. Clones from  the
                     primary library were randomly selected for single pass
                     sequencing."
BASE COUNT       49 a     40 c     57 g     59 t      5 others
ORIGIN


alignment_scores:
            Quality:  126.00             Length:     39
              Ratio:  3.818                Gaps:      2
 Percent Similarity:  84.615   Percent Identity: 76.923

alignment_block:
 US-09-508-832-6 x BF172831   ..

 Align seg 1/1  to: BF172831  from: 1  to: 210

     128 AlaSerIleArgGlnSerGlnGluGluProGluAspLeuArgProGluIl 144
         ||||||:::|||||||||       ||||||   |||:::||||||||||
      95 GCTTCCATGAGGCAGTCT......GAACCTGCAGATATGCGCCCAGAGAT 138

     144 eArgIleAlaGlnGluLeu.ArgArgIleGlyAspGluPheAsnGluThr 160
         |   ||||||||||||||| ||||||||||||||||||||||   :::
     139 ATGGATCGCCCAAGAGTTGGCGGCGTATTGGAGACGAGTTTAACGCTAAC 188

     161 TyrThrArgArgVal 165
         ||||||||||||||
     189 TATACAAGGAGGGTA 203

seq_name: gb_est1:AA629050

seq_documentation_block:
LOCUS       AA629050      501 bp    mRNA            EST       16-OCT-1997
DEFINITION  zu84a06.s1 Soares_testis_NHT Homo sapiens cDNA clone IMAGE:744658
            3', mRNA sequence.
ACCESSION   AA629050
VERSION     AA629050.1  GI:2541437
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 501)
  AUTHORS   Hillier,L., Allen,M., Bowles,L., Dubuque,T., Geisel,G., Jost,S.,
            Krizman,D., Kucaba,T., Lacy,M., Le,N., Lennon,G., Marra,M., Martin
            ,J., Moore,B., Schellenberg,K., Steptoe,M., Tan,F., Theising,B.,
            White,Y., Wylie,T., Waterston,R. and Wilson,R.
  TITLE     WashU-NCI human EST Project
  JOURNAL   Unpublished (1997)
COMMENT     Contact: Wilson RK
            Washington University School of Medicine
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: est@watson.wustl.edu
            This clone is available royalty-free through LLNL ; contact the
            IMAGE Consortium (info@image.llnl.gov) for further information.
            Seq primer: -40m13 fwd. ET from Amersham
            High quality sequence stop: 474.
FEATURES             Location/Qualifiers
     source          1. .501
                     /organism="Homo sapiens"
                     /db_xref="GDB:5932318"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:744658"
                     /clone_lib="Soares_testis_NHT"
                     /sex="male"
                     /lab_host="DH10B"
                     /note="Vector: pT7T3D-Pac (Pharmacia) with a modified
                     polylinker; Site_1: Not I; Site_2: Eco RI; 1st strand cDNA
                     was prepared from mRNA obtained from Clontech Laboratories
                     , Inc., and primed with  a Not I - oligo(dT) primer [5'
                     TGTTACCAATCTGAAGTGGGAGCGGCCGCCCAATTTTTTTTTTTTTTTT 3']. 
                     Double-stranded cDNA was ligated to Eco RI adaptors
                     (Pharmacia), digested with Not I and cloned into the Not I
                     and Eco RI sites of the modified pT7T3 vector. Library
                     went through one round of normalization to Cot5, and was
                     constructed  by Bento Soares and M. Fatima Bonaldo. "
BASE COUNT      155 a    111 c     98 g    137 t
ORIGIN


alignment_scores:
            Quality:  123.00             Length:     31
              Ratio:  4.556                Gaps:      0
 Percent Similarity:  87.097   Percent Identity: 70.968

alignment_block:
 US-09-508-832-6 x AA629050/rev   ..

 Align seg 1/1  to reverse of: AA629050  from: 1  to: 501

     166 PheAlaAsnAspTyrArgGluAlaGluAspHisProGlnMetValIleLe 182
         |||   |||:::|||:::   ||||||||||||||||:::||||||||||
     499 TTTTTGAATAATTACCAAGCAGCCGAAGACCACCCACGAATGGTTATCTT 450

     182 uGlnLeuLeuArgPheIlePheArgLeuValTrpArgArgHis 196
         |:::|||||||||:::|||   |||||||||||||||   |||
     449 ACGACTGTTACGTTACATTGTCCGCCTGGTGTGGAGAATGCAT 407

seq_name: gb_est2:BF259468

seq_documentation_block:
LOCUS       BF259468      768 bp    mRNA            EST       23-FEB-2001
DEFINITION  HVSMEf0019D03f Hordeum vulgare seedling root EST library HVcDNA0007
            (etiolated and unstressed) Hordeum vulgare cDNA clone
            HVSMEf0019D03f, mRNA sequence.
ACCESSION   BF259468
VERSION     BF259468.2  GI:13119934
KEYWORDS    EST.
SOURCE      barley.
  ORGANISM  Hordeum vulgare
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooideae
            ; Triticeae; Hordeum.
REFERENCE   1  (bases 1 to 768)
  AUTHORS   Wing,R., Close,T.J., Kleinhofs,A., Wise,R., Begum,D., Frisch,D., Yu
            ,Y., Anderson,H., Dale,J., Henry,D., Kernodle,S., Palmer,M., Rambo
            ,T., Saski,C., Schwartzbeck,J., Simmons,J., Choi,D.W., Main,D. and
            Wood,T.
  TITLE     Development of a genetically and physically anchored EST resource
            for barley genomics
  JOURNAL   Unpublished (2000)
```

```
COMMENT     On Nov 16, 2000 this sequence version replaced gi:11188497.
            Contact: Wing RA
            Clemson University Genomics Institute
            Clemson University
            100 Jordan Hall, Clemson, SC 29634, USA
            Tel: 864 656 7288
            Fax: 864 656 4293
            Email: rwing@clemson.edu
            Seq primer: AATTAACCCTCACTAAAGGG
            High quality sequence stop: 747.
FEATURES             Location/Qualifiers
     source          1. .768
                     /organism="Hordeum vulgare"
                     /cultivar="Morex"
                     /db_xref="taxon:4513"
                     /clone="HVSMEf0019D03f"
                     /clone_lib="Hordeum vulgare seedling root EST library
                     HVcDNA0007 (etiolated and unstressed)"
                     /tissue_type="Seedling root"
                     /lab_host="TJC121"
                     /note="Vector: lambdaZAP; Site_1: EcoR1; Site_2: Xho1; For
                     more details on library preparation and sequence analysis
                     see http://www.genome.clemson.edu/projects/barley/ To
                     order a clone see http://www.genome.clemson.edu/orders"
BASE COUNT       98 a    322 c    178 g    167 t      3 others
ORIGIN


alignment_scores:
          Quality:  118.50            Length:    134
            Ratio:  1.852               Gaps:      5
 Percent Similarity:  47.761   Percent Identity: 32.836

alignment_block:
 US-09-508-832-6 x BF259468   ..

 Align seg 1/1  to: BF259468  from: 1  to: 768

      4 GlnProSerAspValSerSerGluCysAspArgGluGlyGlyGlnLeuGl 20
        ||||||:::   |||:::   :::   |||
    149 CAGCCCACCGGCGTCACCCTATCCCCGGACCCG..............GC 183

     20 nProAlaGluArgProProGlnLeuArgProGlyAlaProThrSerLeuG 37
         |||:::::::::||||||:::      |||:::   |||   |||
    184 GCCTNTCCGNCACCCGCCTTCGCGCCTNCCCGCCCCTCCACCTTCTGCGC 233

     37 lnThrGluProGlnGlyAsnProAspGlyGluGlyAspArgCysProHis 53
              |||                                 ||||||
    234 CGCCAACCCCC............................TGCCCTTCA 253

     54 GlySerProGlnGlyProLeuAlaPro............ProAlaSerPr 66
                   ||||||:::|||           |||:::|||||
    254 TGTCCTTCGACATACCCGCTCAGTCCAACTCCACGCCGCCCATCTTCACC 303

     66 oGlyProPheAlaThrArgSerProLeuPheIlePheValArgArgSerS 83
        |||||||   |||:::      |||              ||||||||||
    304 GGGCCCATCGGCGGCTCCGGTGCCTCCTTCGAAGACGAGCCGCCGCTCCT 353

     83 erLeuLeuSerArgSer............SerSerGlyTyrPheSerPhe 95
        ||      |||   |||            |||:::|||      |||
    354 CGAGGAGCTCGGCATCAACACGCGCCAGATCTGGCGGAAGACGCTCTCCA 403

     96 AspThrAspArgSerProAlaProMetSerCysAspLysSerThrGlnTh 112
        ::::::   ||||||      |||           :::||||||||||||
    404 TCCTCCACCCGCTCCGCTCCGCCGACCCGT......CGCTCCACGCAGAC 447

    112 rProSerProProCysGlnAlaPheAsnHisTyrLeuSerAlaMetAlaS 129
        ||||      |||:::   :::   :::   |||   |||:::   :::|
    448 GCCGATCTCTCCGGCCCCTTCCTGTTCCTCTTATCCTTCGGTCTCTTCCA 497

    129 er 129
        ||
    498 GC 499
```

```
seq_name: gb_htc:BC007683

seq_documentation_block:
LOCUS       BC007683     1350 bp    mRNA            HTC       12-JUL-2001
DEFINITION  Homo sapiens, postmeiotic segregation increased 2-like 9, clone
            IMAGE:3637150, mRNA.
ACCESSION   BC007683
VERSION     BC007683.1  GI:14712737
KEYWORDS    .
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 1350)
  AUTHORS   Strausberg,R.
  TITLE     Direct Submission
  JOURNAL   Submitted (11-MAY-2001) National Institutes of Health, Mammalian
            Gene Collection (MGC), Cancer Genomics Office, National Cancer
            Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
            USA
  REMARK    NIH-MGC Project URL: http://mgc.nci.nih.gov
COMMENT     Contact: MGC help desk
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: ATCC
            cDNA Library Preparation: Rubin Laboratory
            cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
            DNA Sequencing by: National Institutes of Health Intramural
            Sequencing Center (NISC),
            Gaithersburg, Maryland;
            Web site:       http://www.nisc.nih.gov/
            Contact:        nisc_mgc@nhgri.nih.gov
            Shevchenko,Y., Wetherby,K.D., Beckstrom-Sternberg,S.M.,
            Benjamin,B., Blakesley,R.W., Bouffard,G.G., Brinkley,C., Brooks,S.,
            Dietrich,N.L., Guan,X., Gupta,J., Ho,S.-L., Karlins,E., Legaspi,R.,
            Lim,M., Maduro,Q.L., Masiello,C., Mastrian,S.D., McCloskey,J.C.,
            McDowell,J., Pearson,R., Snyder,B., Stantripop,S., Thomas,P.J.,
            Tiongson,E.E., Touchman,J.W., Tsurgeon,C., Vogt,J.L., Walker,M.A.,
            Zhang,L.-H. and Green,E.D.

            Clone distribution: MGC clone distribution information can be found
            through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov
            Series: IRAL Plate: 12 Row: f Column: 16
            This clone was selected for full length sequencing because it
            passed the following selection criteria: matched mRNA gi: 3327049
            This clone has the following problem: incomplete processing.
FEATURES             Location/Qualifiers
     source          1. .1350
                     /organism="Homo sapiens"
                     /db_xref="LocusID:5387"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:3637150"
                     /tissue_type="Placenta, choriocarcinoma"
                     /clone_lib="NIH_MGC_21"
                     /lab_host="DH10B-R"
                     /note="Vector: pOTB7"
BASE COUNT      333 a    326 c    416 g    275 t
ORIGIN


alignment_scores:
          Quality:  115.00            Length:    187
            Ratio:  1.127               Gaps:     12
 Percent Similarity:  54.545   Percent Identity: 32.086

alignment_block:
 US-09-508-832-6 x BC007683   ..

 Align seg 1/1  to: BC007683  from: 1  to: 1350

     14 ArgGluGlyGlyGlnLeuGlnProAlaGluArgProProGlnLeuArgPr 30
        |||:::||||||      :::   |||      |||            ||
     98 CGAAGTGGCGGCGGAGGCGAAGGGGCGGCGGGACCC.........GGGCC 138
```

```
 30 oGlyAlaProThrSerLeuGln......ThrGluProGln......GlyA 43
    ||||      :::|||||||          |||:::|||:::      |||:
139 TGGCCCGTATGTGTCCTTGGCGGCCTAGACTAGGCCGTCGCTGTATGGTG 188

 43 snProAspGlyGluGlyAspArgCysProHisGlySerProGlnGlyPro 59
    ::|||:::|||   |||:::      |||:::|||   |||   |||
189 AGCCCCAGGGAGGCGGATCTGGGCCCCCAGAAGGACACCCGCCTGGA... 235

 60 LeuAlaProProAlaSerProGlyProPheAlaThrArgSerProLeuPh 76
    :::||||||   |||   |||||||||:::::::::|||:::   |||
236 TTTGCCCCGTAGGCCCGGCCCGGGCCCCTCGGGAGCAGAACAGCCTTG.. 283

 76 eIlePheValArg.ArgSerSerLeuLeuSerArgSerSerSerGlyTyr 92
          ||||||    ::::::      ||||||:::::::::
284 .......GTGAGGTGGACAGGAGGGGACCTCGCGAGCAGACGCGCGC... 323

 93 PheSerPheAspThrAspArgSerProAlaProMetSerCysAspLysSe 109
             :::|||   ||||||||||||   |||   :::   ::
324 .........GCCAGCGACAGCAGCCCCGCCCCGGCCTCTCGGGAGCCGGG 364

109 rThrGlnThrProSerProProCysGlnAlaPheAsnHisTyrLeuSer. 125
    :   :::      :::   |||      ::::::::::|||   |||
365 GGGCAGAGGCTGCGGAGCCCCAGGAGGGTCTATCAGCCACAGTCTCTGCA 414

126 ..............................AlaMetAlaSerIleArg 131
                                   |||   ||||||:::
415 TGTTTCCAAGAGCAACAGGAAATGAACACATTGCAGGGGCCAGTGTCATT 464

132 Gln.Ser.................GlnGluGluProGluAspLeuArgP 142
    ||| :::                 ||||||||   ::::::|||   |
465 CAAAGATGTGGCTGTGGATTTCACCCAGGAGGAGTGGCGGCAACTGGACC 514

142 roGluIleArgIleAlaGlnGluLeuArgArgIleGlyAspGluPheAsn 158
    ||:::   :::||||||                  ||||||   :::
515 CTGATGAGAAGATAGCA..............TACGGGGATGTGATGTTG 549

159 GluThrTyrThrArgArgValPheAlaAsn...AspTyrArgGluAlaGl 174
    |||:::|||::::::   |||   ::::::   ||||||::::::|||::
550 GAGAACTACAGCCATCTAGTTTCTGTGGGGTATGATTATCACCAAGCCAA 599

174 uAspHis 176
    :   |||
600 ACATCAT 606
```